

EV(FRP51)-ETA FUS1
Monoclonal antibody

ALIGNMENTS

Fv(FRP51)-ETA fused
Monoclonal antibody

RESULT

IDENTITY NUM

NY
OC
Syr+hot+ic
pro-drug conv

PN WO9734634-A1.
DE 35 000 1007

PR 20-MAR-1996; US 013703.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
PI Cheung NV, Guo H, Larson SM, Rivlin K, Sadelain M:
PF 20-MAR-1997; 004427.

WT, 2147550/44
Recombinant single
PT

PT	useful to detect tumour cells expressing GD2 and to target
PT	therapeutic agents, e.g. toxins, to such cells
PS	Disclosure; Page 11; 31pp; English.
CC	The present sequence encodes a recombinant single chain peptide,
CC	5F11-scFv. The peptide is an antibody construct comprising the variable
CC	regions of the heavy and light chains of an antibody against
CC	disialanglioside (GD2) as a single chain Fv fragment (scFv). GD2 occur
CC	in many tumours types including neuroblastoma, osteosarcomas and other
CC	soft tissue sarcomas, medulloblastomas, high grade astrocytomas,
CC	melanomas and small cell lung cancer. The peptide can be detectably
CC	labelled, preferably with 99m-Tc, for tissue imaging of cells expressing
CC	GD2. It can also be used to target delivery of a therapeutic or
CC	pre-therapeutic agent, such as a toxin, streptavidin or a pro-drug
CC	converting enzyme, to cells expressing GD2. The peptide may further
CC	comprise CD8 to facilitate the formation of GD2-targeted lymphocytes.
CC	T cells containing the peptide can also be used to target GD2-producing
CC	tumour cells.
SQ	Sequence 717 BP; 180 A; 181 C; 195 G; 159 T;
Query Match 99.7%; Score 715; DB 1; Length 717;	
Best Local Similarity 100.0%; Pred. No. 5.2e-196;	
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps	
QY	1 caggtagaactgcagcagtcaggaccctgaactcgtgtnagcctggggcttcagtgaagata 60
DB	1 CAGGTGAAACTGCAGCAGTCAGGACCTGAACCTGGTGNAGCCTGGGGCTTCAGTGAAGATA 60
QY	61 tctctgcaagactcctgganacaaattcactgaatacaccatgcactcgggtgaagcagac 120
DB	61 TCTCTGCAAGACTTCTGGANACAAATTCATGTAATACCCATGCACCTGGGTGAAGCAGAC 120
QY	121 catggaaagagccttgagtgagtgaggattgaattaatcctaaacaatggtggtactaacac 180
DB	121 CATGGAAAGAGCCTTGAGTGGATTGGAGCTATTATCTCTAACAAATGGTGTACTAACTAC 180
QY	181 aagcagaagttaagggcgaagccacattgactgtagacaaagtctccagcacagcctac 240
DB	181 AAGCAGAAGTTTAAAGGGCAAGGCCACATTGACTGTAGACAAGTCTCTCCASCACAGCTAC 240
QY	241 atggagcttcgacagctgcacattcaggattctcgagttcttactgtgcagaagatact 300
DB	241 ATGGAGCTTCGACAGCTGCACATCTGAGGATCTGCGAGTCTATTACTGTGCAAGAGATACT 300
QY	301 acggtcccggttgcttactgggttcccaaggagacacaggtcaccgctccctcaggttgagac 360
DB	301 ACGGTCCCGGTTGCTTACTGGGTTCCAAGGACACACAGGTCAACGCTCTCTTCACGTTGAGGC 360

DR WPI: 99-243623/20.
DR P-SDB: Y05363.
PT Single-stranded antibody against hepatitis B virus core protein,
PT applicable as (gene) therapeutic agents for treatment of hepatitis B
PT viral infections
PS Claim 5: Page 55-57: 72pp; English.
CC This sequence encodes the single-stranded antibody of the invention, that
CC has the capability of binding to a hepatitis B virus (HBV) core protein.
CC Therapeutic agents can be formulated with the antibody for treatment of
CC HBV infections by stopping proliferation of the virus through inhibition
CC of viral DNA synthesis, and the gene encoding the antibody can be applied
CC as an agent to gene therapy.
SQ Sequence 894 BP: 217 A; 225 C; 237 G; 215 T;

Query Match 73.1%; Score 523.8; DB 1; Length 894;
Best Local Similarity 85.6%; Pred. No. 4e-141;
Matches 524; Conservative 0; Mismatches 84; Indels 21; Gaps 3

Qy	1	caagtgaaactgcagcagtcagacacctggaactggtgnagcctggggcttcagtgaagata	60
Db	118	CAAGTGAAGCTGCAGGAGTGCAGACCTGAGCTGGAGAACCTTCGCGCTTCAGTGAAGATA	177
Qy	61	tcctgaaagactctgaganacaattcactgaatacacaccatgcactgggtgaagcagac	120
Db	178	TCTTGCNAGGCTTCTGGTTACTCATTTCACTGGCTACAACTGAAGTGGGTGAACAGAC	237
Qy	121	catggaagagcccttgagtgattgaggtattaatcctaacaatggtgggtactaaact	180
Db	238	AATGGAAAGAGCCTTGAGTGGATTGGATATATTATTCCTTACAATGGTGGTACTGGCTAC	297
Qy	181	aagcagaagttcaaggccaagggccaatgcatgtgataagacaagtctctcagcacagcctac	240
Db	298	AACGAGAAGTTCAAGACGAAGGCCACATTGACTGTAGCAAAATCTCCAGCACAGCCTAC	357
Qy	241	atggagctccgagcctgcacatcagagattctgcagttattactgttgcgaagatact	300
Db	358	ATGCACTGAGCAGCCTGACATCTGAGGACTCTGCAGTCTATTACTTGTGCAGA-----	412
Qy	301	acggtcccgcttctactgggtgccaaaggacacaggtcacccgtctccctcaggtggagcg	360
Db	412	---CTGGACTTGACTACTTGGGGCCAAAGGACCACCGTCAACCGTCTCCTCAGGTGGAGGC	468
Qy	361	ggttcaggcggaggtggctctggcgtggcgatcggacatcgagctcactcagttctcca	420
Db	469	GGTTACGGCGGAGGTGGCTCTTGGCGGTGGCGATCGGACATCGAGTCTACTCAGTCTCCA	528
Qy	421	gcactatctctcatctccaggggagaaggtccaccatgacctgagttggcagctcaagt	480
Db	529	ACCACATGCTGCATCTCCGGGGAGAAGATCACTATCACTTGCAGTCCGAGCTCAAGT	588
Qy	481	ataag-----ttacatgcactggtaccagcagaagcctgtcacctccccccaaaagatgg	534
Db	589	ATAAGTTCCAATTACTTGCAATGGTATCAGCAGAAGCCAGGATTCCTCCCTTAACCTCTTG	648
Qy	535	attatgacacatccaaactggctcttgagtgccctgtcgtctcagttggcagttgggtct	594
Db	649	ATTTATAGGACATCCAATCTGGCTCTCGAAATCCCAAGTCCGTTTCAGTGGCAGTGGGTCT	708
Qy	595	ggagactcttattctccaaatcagcacatgagagcttagatctgacctctccactattac	654
Db	709	GGGACCTCTTACTCTCACAATTGGCACCATGGAGGCTGAAGATGTTGCCACTTTACTAC	768
Qy	655	tgcacatcagcggagtagttaccgc-----tcacgttcgggtgctgggacacagttggaa	708
Db	769	TGCCAGCAGGGTAGTAGTATACCACCAATATTCAAGTTCGGTCTCGGTCGGGACAAAGTTGGAA	828
Qy	709	ataaacagg	717
Db	829	ATAAAACGG	837

RESULT

T04023	ID	T04023 standard; cDNA; 726 BP.
AC	AC	T04023;
DT	DT	02-JUL-1996 (first entry)
DE	DE	Anti-EGFR single chain antibody (Clone 5 F 1).
KW	KW	Single chain antibody; antibody; epidermal growth factor receptor;
KW	KW	EGFR; tumour; cancer; glioma; melanoma; carcinoma; diagnosis;
KW	KW	assessment; phage antibody library; ss.
OS	OS	Mus musculus.
FS	FS	Location/Qualifiers
FT	FT	1. .726
FT	FT	/*tag= a
FT	FT	/product= Single-chain Fv, heavy and light chain
FT	FT	plus linker.
PN	PN	W09525167-A1.
PD	PD	21-SEP-1995.
PF	PF	16-MAR-1995; E00978.
PF	PF	17-MAR-1994; EP-104160.
PR	PR	02-DEC-1994; EP-118970.
PA	PA	(MERE.) MERCK PATENT GMBH.
PI	PI	Adan J, Ansell KH, Bendig MM, Blasco F, Guessow D;
PI	PI	Kettleborough AC, Mitjans F, Piuats J, Rosell E;
DR	DR	WPI; 95-336972/43.
DR	DR	P-PSDB; R79870.
PT	PT	Anti-EGFR antibodies and single chain Fv antibody fragments -
PT	PT	obtained from phage-antibody libraries, useful for diagnosis and
PT	PT	therapy of tumours
PS	PS	Claim 4; Page 66-67; 93pp; English.
CC	CC	Anti-epidermal growth factor receptor (EGFR) single chain antibodies
CC	CC	and antibodies constructed from anti-EGFR antibody fragments can be
CC	CC	used for diagnosis of tumours and assessment of tumour growth in
CC	CC	vitro and in vivo. They may also be used in a pharmaceutical
CC	CC	composition for the therapy of e.g. melanomas, gliomas and carcinomas.
CC	CC	The antibodies and fragments are derived from mice but are humanised
CC	CC	so as to cause minimum reaction against them. They are produced
CC	CC	using the phage antibody library. (See T04011-T04026 and
CC	CC	R79858-R79873)
Sequence	Sequence	726 BP; 174 A; 192 C; 206 G; 154T;

Query Match	72.7%	Score	521.6;	DB 1:	Length	726;			
Best Local Similarity	84.4%;	Pred. No.	1.6e-140;						
Matches	613;	Conservative	0;	Mismatches	101;	Indels	12;	Gaps	
Qy	1	cagtgaaactcgacgcagctcgagcagctgaactgctgnagcctggggcttcagtgaaagata	60						
Db	1	CAGTGAAACTCGAGAGNGCTGGGGCTGAACCTGTGAACCTGGGGCTTCAGTGAAGTTG	60						
Qy	61	tccctgcaagactcttctggnacaattcactgaaatacacccatcactcgtgtgaagcagc	120						
Db	61	TCCTGCAAGGCTTCCGGCTACACTTCACGAGCACTGGATGCATGTGGTGAAGCAGAGG	120						
Qy	121	catggaagagccttcagtgattgaggtattaatccttaacaatggtggtactaactac	180						
Db	121	CTTGGACAAGCCCTTGATGGATCGGAGAGATTAAATCCAGACGGCGCTACTAACTAC	180						
Qy	181	aagcagaagtccaaggggcaaggccacattgactgtagacaagtcctccagcacagctac	240						
Db	181	AATGAGAAATTCAGAGCAAGGCCACACTGACTGTAGACAAATCTCCAGCACAGCCTAC	240						
Qy	241	atgaggtccgcagcctgacatctgaggtctcgagttctgagctattactgtgcaag-agatac	299						
Db	241	ATGCAACTCAGACGCTGCATCTGAGGACTCTGCGGTCTATTACTGTGCCAGTCGGGAC	300						
Qy	300	tacggtcccg-----tttgctactgggtccaaaggaccacggtcacccgtctcc	348						
Db	301	TATGATTATCACGCGAGCGGTACTTTGACTACTTGGGGCCAAAGGGACAACGGTACCGTCTCC	360						
Qy	349	tcaggtgagcggttcaggcgaggtgctctggcggtggcgatcgacatcagctc	408						
Db	361	TCAGTGGCGGTGCTCGGGCGGTGGTGGGTTCGGGTGGCGGGATCTGACATTGAGCTC	420						
Qy	409	actcagttcccgacaatcatgtctgcatctccagggggagaaggtcaccatgacctgcagt	468						

CC used in the antibody-firefly luciferase fusion protein of the invention
CC in which an antibody part consisting of a peptide having antibody
CC activity is combined with an enzyme part consisting of firefly
CC luciferase.
SQ Sequence 732 BP; 179 A; 183 C; 212 G; 158 T

Query Match	72.6%	Score 520.4;	DB 1;	Length 732;
Best Local Similarity	84.5%;	Pred. No. 3.6e-140;		
Matches 610;	Conservative	0;	Mismatches 103;	Indels 9;
				Gaps 2;

[illegible]

Qy	182	agcagaagtccaaggc	caagccacattg	ctactgtagac	agctctccagc	cacacgctctaca	241.1
Db	188	ATGGAAGTTCAAGGGTAAAGTACATG	CTGCTGAGACAAATCCTTC	CAGCACACGCCACA			247.7
Qy	242	tggagctccgcagcctg	acatctggagat	ctcgactctattact	gtgcgaagagata	lacta	301.5
Db	248	TGACGCTCAGCAGCCTAACATCTG	AAGACACTCTGGGGTCTATTTCT	GTACAAAGA---ATGG			304.3

Qy	302	cggtcccgattgtctactagggtccaaaggaccacggtcaccggtctctcaggtggaggcg	361
Db	305	GGGTACACACGGGACTACTTGGGGCCAAAGGGACACAGGTTCACCGTCTCCTCAGGTGGAGCG	364
Qy	362	gttcaggcggaagttgggtctcggggtggcggaatcggaatcgagctcaactcagctctccag	421
Db	365	GTTCACGCGGAGGTGGCTCTGGCGTGGCGGATCGGAATCGGAATCAGCTCTCCAG	424

[illegible]

Qy	536	tttatgacacatcaaaactggctctctgagtcgctcttcagtcgagcagtcgggctcg	5395
Db	545	TTTATGGCACATCCAACTGSGCTCTGGAAATCCCTGTTCCGTTTCAGTGGCAGTGATCTG	604
Qy	596	ggacctcttatctctctcaaatcagcagcagctgagcgctgtagatgctgcacacttatct	655

QY	656	gccatcagcgagtagttaccgctcacgttcggtctgggcacacagtctggaataaac	7155
Db	665	GCCAGCAGTGGAGTAGTAACCCATTTCACGTTCTGGCTCGGGGCCAACGCTGGAATAAAAC	724
QY	716	gg 717	
Db	725	GG 726	

RESULT	7
V23580	
ID	V23580 standard; cDNA to mRNA; 2364 BP.
AC	V23580;
DT	16-JUL-1998 (first entry)
DE	Antibody-Firefly luciferase fusion protein gene.
KW	Firefly luciferase; antibody-luciferase fusion protein; ds.
OS	Luciola cruciata.
OS	Luciola cruciata.
Key	Location/Qualifiers
FH	key

FT	1. .2364	
FT	/*tag= a	
FT	/transl_except= (pos: 670, aa: Glu)	
FT	/transl_except= (pos: 739, aa: Trp)	
FT	/transl_except= (pos: 1369, aa: Ala)	
FT	/note= "no stop codon given"	
FT		
FN	J09187281-A.	
PD	22-JUL-1997.	
PF	09-JAN-1996; 001812.	
PR	09-JAN-1996; JP-001812.	
PA	(KIKK) KIKKOMAN CORP.	
DR	WPI; 98-275089/25.	
DR	P-PSDB; W53882.	
PT	Antibody-firefly luciferase fused protein - and related products	
PT	i.e. firefly luciferase fused gene, recombinant DNA and its	
PT	preparation	
PS	Disclosure: Page 13: 17pp; Japanese.	
CC	This sequence encodes the fusion protein of the invention. The protein	
CC	a antibody-firefly luciferase fusion protein, in which an antibody para	
CC	consisting of a peptide having antibody activity is combined with an	
CC	enzyme part consisting of firefly luciferase.	
CC	Sequence 2364 BP: 697 A: 444 C: 559 G: 664 T:	
SQ		

Query Match 72.2%; Score 517.8; DB 1; Length 2364;
Best Local Similarity 84.3%; Pred. NO. 3e-139;
Matches 608; Conservative 0; Mismatches 104; Indels 9; Gaps

[illegible]

181
122 atggaagagccttgadtgattgagggtattaatcctaacaatgggtggtactaataca 181
123 |
124 |
125 |
126 |
127 |
128 CTGACACAGGCTCTTACGTGATTGGACAGATTATCTGGAGATGGTGATACGAATTAC 187
129 |
130 |
131 |
132 |
133 |
134 |
135 |
136 |
137 |
138 |
139 |
140 |
141 |
142 |
143 |
144 |
145 |
146 |
147 |
148 |
149 |
150 |
151 |
152 |
153 |
154 |
155 |
156 |
157 |
158 |
159 |
160 |
161 |
162 |
163 |
164 |
165 |
166 |
167 |
168 |
169 |
170 |
171 |
172 |
173 |
174 |
175 |
176 |
177 |
178 |
179 |
180 |
181 |
182 agcagaagtccaaggcgaaggccacattgactatgacaagtcctccagcagagcctaca 241
183 |
184 |
185 |
186 |
187 |
188 ATGCAAAAGATTCAAGGGTAAAGTATCATCTGATCTGCAGACAATCTCTCCAGTACAGCCCA 247
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 |
242 |
243 |
244 |
245 |
246 |
247 |
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 |
302 |
303 |
304 |
305 |
306 |
307 |
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 |
362 |
363 |
364 |
365 |
366 |
367 |
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 |
422 |
423 |
424 |
425 |
426 |
427 |
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 |
482 |
483 |
484 |
485 |
486 |
487 |
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 |
542 |
543 |
544 |
545 |
546 |
547 |
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 |
602 |
603 |
604 |
605 |
606 |
607 |
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 |
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
68

Qy	242	tggagctccgcagcctgacatctgaggaattctcagctctattactgtgcagagatacta	301
Db	248	TGCAGCTCAGCAGCCTAACATCTGAAGACTCTCGGGTCTATTCTGTACAAAGA--ATGG	304
Qy	302	cggtcccgcttgccttactgggtccaaggagccacggtcacccgtctcctcaggtggaggcg	361
Qy	305	cggtcccgcttgccttactgggtccaaggagccacggtcacccgtctcctcaggtggaggcg	364

Y	362	gttcaggcggaagtggctctgggcggtgagcgatcgacatcgagctcaactcagctctccag	421
Db	365	gttcaggcggaagtggctctgggcggtgagcgatcgacatcgagctcaactcagctctccag	424
Y	422	caatcatgtctgcatactccaggggggaaggttcaccatgcagctgcagtcggcagctcaagta	481

Qy	482	taagtt-----acatgacgtgtaccagcagagaagcgtctcaacctcccacaaagatgga	535
Db	485	TAAgTTCAGCAACTTTGCATCTGGTACCAGCAGAAAGTCAGAAACTCCCCCAAACTCTGGA	544
Qy	536	ttaaTgacacatccaaactggctcttgagctccctgctgcttcagtggcagtggtgctg	595

[illegible]

PT	therapy of tumours			
PS	Claim 4; page 68-69; 93pp: English.			
CC	Anti-epidermal growth factor receptor (EGFR) single chain antibodies			
CC	and antibodies constructed from anti-EGFR antibody fragments can be			
CC	used for diagnosis of tumours and assessment of tumour growth in			
CC	vitro and in vivo. They may also be used in a pharmaceutical			
CC	composition for the therapy of e.g. melanomas, gliomas and carcinomas			
CC	The antibodies and fragments are derived from mice but are humanised			
CC	so as to cause minimum reaction against them. They are produced			
CC	using the phage antibody library. (See T04011-T04026 and			
CC	R79858-R79873)			
SQ	Sequence 726 BP: 171 A: 193 C: 205 G: 157 T:			

Query Match 71.5%; score 512.6; DB 1; Length 726;
Best Local Similarity 83.7%; Pred. No. 6.1e-138;
Matches 607; Conservative 0; Mismatches 106; Indels 12; Gaps

QY	2	aggtgaactcgagcagtcaggacctgaactgctgnaagcctgggggttcagtgaaagatat	61
Db			
Db	2	AGGTCAAAGCTGCAGCAGTCAGGGGCTGAACGTGTGAAGCTGGGGCTTCAGTGAAGTTGT	61
QY	62	ctgtcaagactcttganaacaaattcaactgaatacacaccatgcactgggtgaagcagaagcc	121
Db			
Db	62	CTGCAAGGCTTCGGGTACACCTTCACAGCCACTTGGATCACTGGGTGAAGCAGAGGG	121
QY	122	atggaagaagccttgagtgatggaggatttaatactcaacaatggtggtactaaactaca	181
Db			
Db	122	CTTGGCAAGGCTTGAGTGGATCGGACATTTAATCCACCAACGGCGCTACTACTACTA	181
QY	182	agcagaagttcaagggaagggccacattgactgtagacaagtcctccacacagcctaca	241
Db			
Db	182	ATGAGAAATTCAGAGCAAGGCCACACTGACTGTAGACAAATCCTCCAGCAGCGCTACA	241
QY	242	tggagctcgcagccctgacatctggaagattctgcagttctattactgtgcaag-agatact	300
Db			
Db	242	TCGAATCTACAGCCCTGACATCTGAGGACTGCTCGGTCTATTACTGTGCAGTCGGGACT	301
QY	301	acggtcccg-----tttgcttactgggtccaaaggacacaggtccacggtctcct	349
Db			
Db	302	ATGATTACGACGGACGGTACTTTTGACTACTGGGGCCCAAGGGACCGGTCAACGTCCT	361
QY	350	caggtggaggcgttcaggcggagggtggctctgctggcgtggcggatcggacatcgagctca	409
Db			
Db	362	CAGGTGGCGGTGGCTCGGGCGGTGGTGGGTGGGTGGCGCGGATGTGACATTGAGCTCA	421
QY	410	ctcagttccagcaaatcatgtctgcattctccaggggagaaggtccacatgacctgcagtg	469
Db			
Db	422	CCAGTCTCCAAACAATCATGCTGTGCATCTCCAGGGAGAAGGTACCATGACTTGCAGTG	481
QY	470	gcagctcaagataaagttacatgcactggtaccagcagaagcctgtccactcccccaaa	529
Db			
Db	482	ACAGCTCAAGTGTAAATTACATGCTACTGTACCAGCAGAAAGCAGAGATCCTCCCCAGAC	541
QY	530	gatggatttatgacacatccaaactggcctctgtgagtcctcgtcgtctcagtgggcagtg	589
Db			
Db	542	TTCTGATTTATGACATCCAACTGGCTTCTGGAGTCCGCTTTCGCTTCAGTGGCAGTG	601
QY	590	ggtctgggaaccttattctctcaaatcagcagaatggaggctgtagatgctgcaactt	649
Db			
Db	602	GGTCTGGGACCTTCTCTCTCACAAATCAGCCGAATGGAGGCTGAAGATGTGTCACATT	661
QY	650	attactggcatcgagcgagtagttaccgctcagttctcggtgctgggacacagtttgaaa	709
Db			
Db	662	ATTACTGCCAGCAGTGGAGTAGTTACCGCTACAGTTCGGTCTGGTGGGACCAAGCTGAAA	721
QY	710	taaaa	714
Db			
Db	722	TAATAA	726

RESULT	10
Q90663	

ID	Q90663	standard; DNA; 810 BP.
AC	Q90663;	
DT	26-JAN-1996	(first entry)
DE	MFE-23	antibody coding sequence.
KW	Antibody; MFE-23; carcinoembryonic antigen; CEA; colorectal tumour; therapy; diabody; ds.	
OS	Mus musculus.	
FS	Key	
FT	Location/Qualifiers	
FT	misc_feature	79..438
FT		/*tag= a
FT		/note= "variable heavy chain"
FT	misc_feature	439..483
FT		/*tag= b
FT		/note= "linker region"
FT	misc_feature	484..801
FT		/*tag= c
FT		/note= "variable light chain"
PN	W09515341-A1.	
PD	08-JUN-1995.	
PF	05-DEC-1994; G02658.	
PR	03-DEC-1993; GB-024807.	
PA	(CANC-) CANCER RES CAMPAIGN TECHNOLOGY.	
PI	Begent RJ, Chester KA, Hawkins RE;	
DR	WPI: 95-215234/28.	
DR	P-PSDB: K75719.	
PT	Antibody for carcinoembryonic antigen - for treatment and diagnosis	
PT	of colorectal cancer	
PS	Claim 3; Page 48-49; 72pp; English.	
CC	This sequence represents the coding sequence for the MFE-23 antibody	
CC	The encoded protein is an antibody against carcinoembryonic antigen	
CC	(CEA). CEA is a marker antigen for cancer imaging and therapy. In the	
CC	MFE-23 antibody sequence was obtained using phage technology. In the	
CC	process, mice were immunised with CEA. The antibody variable region	
CC	genes obtained from these mice were then amplified from cDNA and cloned	
CC	as a single chain Fv (scFv) into bacteriophage vectors, producing a	
CC	library. The phages that bound to biotinylated CEA were selected and	
CC	amplified, and this sequence (and the protein it encodes) were selected	
CC	The MFE-23 antibody was found to have good specificity and affinity for	
CC	CEA, meaning that it can be used in targeted anti-tumour therapies. If	
CC	humanised antibody with the complementarity determining regions of MFE	
CC	may be made by CDR grafting. The antibody may be used for the treatment	
CC	by surgery or therapy of a colorectal tumour, or in the diagnosis of	
CC	colorectal tumour. MFE-23 may also be used to make diabodies (bivalent	
CC	or bispecific antibody fragments which bind to two different antigens	
CC	and may be linked to an antitumour agent or a detectable label.	
CC	Sequence	810 BP; 189 A; 220 C; 179 T;

Query Match	70.7%;	Score 507.2;	DB 1;	Length 810;
Best Local Similarity	82.2%;	Pred. No. 2.3e-136;		
Matches 597: Conservative	0;	Mismatches 120;	Indels 9;	Gaps 1;

Qy	1	cagtgaaactgacgacgtcagagacctgaactggtgnagcctgggcttcaagtgaagata	60
Db	79	CAGTGAAACTGCACGAGCTTGGGGCAGAACCTTGTGAGGTCAGGGACCTCAGTCAAGTTG	138
Qy	61	tctgcgaagactctcganacaaattcactgaatacaccatgcac'tgggtggaagcagagc	120
Db	139	TCCTGTGCAGCTTCTGGCTTCAACATTAAAGACTCTATATGACACTGGTTGAGCGAGGG	198
Qy	121	catggaagagacctgagtggtgattggaggttattaatcctcaacaatggtggtactaacac	180
Db	199	CCTGAACAGGGCCCTGGAGTGGATTGGATGGATTGATCCTCAGGAATGGTGACTGAATAT	258
Qy	191	aagcagaagttcaaggcgacagggccacattgactgtgagacaag'tcctccagcacagcctac	240
Db	259	GCCCCGAGTTCAGGCGCAGGGCCACTTTTACTACAGACACATCTCTCCACACAGGCTAC	318
Qy	241	atgaggtcccgagcctgcacatctgaggaattctgcagttctattactgtgcagaagataact	300
Db	319	CTGCAGCTCAGCAGCCCTGACATCTGAGGAGCACTGCCCTCTATTATTGTAATGAGGGGACT	378
Qy	301	-----acggtcccgatttgcattactgcattgcggtccaaaggacccagat'caccgtctctctca	351

Monoclonal antibody
ScFv-BSRNase fusio

ALIGNMENTS

44	274	38.4	336	1	T97441
45	273	38.2	1235	1	Q70663

DT 06-APR-1998 (first entry)
AC 180310;
DE Single chain anti-disialoganglioside GD2 antibody 3G6-scfv.
KW Antibody construct; disialoganglioside; GD2; single chain Fv fragment;
DE scFv; tumour; neuroblastoma; osteosarcoma; soft tissue sarcoma;
KW tissue imaging; target delivery; toxin; streptavidin;
KW pro-drug converting enzyme; GD2-targeted lymphocyte; ss.
OS Synbetic.

Scoring table: IDENTITY_NUC

pro-drug conversion
synthetic

PN WO9/34634-A1.
PD 25-SEP-1997.

PF 20-MAR-1997; 004421.
 PR 20-MAR-1996; US-013703.
 PA (SLOAN) SLOAN KETTERING INST CANCER RES.
 PI Cheung NV, Guo H, Larson SM, Rivlin K, Sadelain M;
 DR WPL: 97-474996/44.
 PT Recombinant single chain anti-disialoganglioside GD2 antibody -
 PT useful to detect tumour cells expressing GD2 and to target
 PT therapeutic agents, e.g. toxins, to such cells
 PS Disclosure: Page 11-12; 31pp; English.

DK
PT
NET, 5/4/33
Recombinant s

	useful to detect tumour cells expressing GD2 and to target
PT	therapeutic agents, e.g. toxins, to such cells
PS	Disclosure: Page 11-12; 31pp: English.
CC	The present sequence encodes a recombinant single chain peptide,
CC	366-scrv. The peptide is an antibody construct comprising the variable
CC	regions of the heavy and light chains of an antibody against
CC	dialoanglioside (GD2) as a single chain Fv fragment (scrV). GD2 occurs
CC	in many tumours types including neuroblastoma, osteosarcomas and other
CC	soft tissue sarcomas, medulloblastomas, high grade astrocytomas,
CC	melanomas and small cell lung cancer. The peptide can be detectably
CC	labelled, preferably with 99m-Tc, for tissue imaging of cells expressing
CC	GD2. It can also be used to target delivery of a therapeutic or
CC	pre-therapeutic agent, such as a toxin, streptavidin or a pro-drug
CC	converting enzyme, to cells expressing GD2. The peptide may further
CC	comprise CD8 to facilitate the formation of GD2-targeted lymphocytes.
CC	T cells containing the peptide can also be used to target GD2-producing
CC	tumour cells.
CC	Sequence 714 BP: 174 A: 168 C: 205 G: 167 T:
SQ	
	Query Match 100.0%; Score 714; DB 1; Length 714;
	Best Local Similarity 100.0%; Pred. No. 1.2e-187;
	Matches 714; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 agtatgtgatgacccagactcccataatcttcgttgtatcatcgaggagacagggttac 60
DB	
DB	1 AGTATTGTGATGCCAGCAGACTCCCAAAATTCCTGTTGTATCAGACGGAGACAGGGTTACC 60
QY	61 ataacctcaaggccagtcagagtgtgagtgaatgatgtggcttggtaccacaagagcca 120
DB	
DB	61 ATAACTCGAAGSCCAGTCAGAGTGATGATGATGATGTTGCCTGGTATCCAACAAGAAGCCA 120
QY	121 gggcagctctccgaacctgcgtatatctctgcattccaatcgctacactggagtcctctgat 180
DB	
DB	121 GGSCAGCTCCGAAACTGCTGATATACTCTGCATCCAAATCGCTACACTGGAGTCCCCTGAT 180
QY	181 cgcttcaactggcagttgatatgggacggatttcaactttcaccatcagcactgtcgaggt 240
DB	
DB	181 CGGTTTCACCTGGCAGTGGATATGGACGGGATTTACATTTCCACCATCAGCACTGTGCGAGGCT 240
QY	241 gaagacctggcagttkattctctgcagcaggattatagctcgtcgaggggggccaaag 300
DB	
DB	241 GAAGACCTGGCAGTTTTATTCTCTGTACAGAGGATTTATAGTCGCTCGAGGGGGGACCAAG 300
QY	301 ctggaaaaataaaggctggagcggttccagcgagggtggctctggcgttgccgattcgcgag 360
DB	
DB	301 CTGGAAATAAAGCTGAGGCGGTTTCAGCGGAGGTGGCTCTGGCGTGTGGCATCGCAG 360

QY 361 gtgcagggtgaaggagctcaggacccctggcctggtggtgcgcgcctcacagagagcctgtccatacact 420
 Db 361 GTGCAGGTTGAAGGAGTCAAGGACCTGGCCCTGGTGGCGCCCTCACAGAGCCTGTCCATCACT 420
 QY 421 tgcactgtctcctgggttttcttaaccattatggtgtacactgggttcgcagccctcca 480
 Db 421 TGCACGTCTCTGGGTTTCTATACCAATATATGGTGTACACTGGGTGCGCAGCCCTCCA 480
 QY 481 ggaagggtcctggagtggtgggagtaataatgggtggtggaagacacaaattataattcg 540
 Db 481 GGAAGGGTCTGGAGTGGCTGGGAGTAATATGGGTGGTGAAGCACAAATATATAATTCG 540
 QY 541 gctcttatgtccagactagcatcagcaagacaaactccagagaccagttttcttaaaa 600
 Db 541 GCTCTTATGTCCAGACTGAGCATCAGCAAGGACAACTCCAAGAGCCAAAGTTTCTTAAAA 600
 QY 601 atgaacagtctgaaactgatacacagccatgtactactgtccagtgcggggggttaac 660
 Db 601 ATGACAGCTGTGCAACCTGTATGACACACCCATGTACTGTGCCAGTCGGGGGGGTAAC 660
 QY 661 tacggctatgctttggactactgggttcaggaaacctcagtcaccgtctctcca 714
 Db 661 TACGGCTATGCTTTGGACTACTGGGTCAAGAACCTCAGTCAGCTCAGCTCTCTCA 714

RESULT 2
 T86312
 ID T86312 standard; DNA; 1173 BP.
 AC T86312;
 DT 06-APR-1998 (first entry)
 DE Single chain anti-disialoganglioside GD2 antibody 3G6-scrFv-streptavidin.
 KW Antibody construct; disialoganglioside; GD2; single chain Fv fragment;
 KW scFv; tumour; neuroblastoma; osteosarcoma; soft tissue sarcoma;
 KW tissue imaging; target delivery; toxin; streptavidin;
 KW pro-drug converting enzyme; GD2-targeted lymphocyte; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 1..714
 FT /*tag= a
 FT /note= "encodes single chain Fv fragment of
 FT antibody 3G6"
 FT misc_feature 715..738
 FT /*tag= b
 FT /note= "linker sequence"
 FT misc_feature 739..740
 FT /*tag= c
 FT /note= "encodes streptavidin"
 FT WO9734634-A1.
 PN 25-SEP-1997.
 PD 20-MAR-1997; U04427.
 PR 20-MAR-1996; US-013703.
 PR (SLOK) SLOAN KETTERING INST CANCER RES.
 PA Cheung NV, Guo H, Larson SM, Rivlin K, Sadelain M;
 PI WPI; 97-479996/44.
 DR Recombinant single chain anti-disialoganglioside GD2 antibody -
 PT useful to detect tumour cells expressing GD2 and to target
 PT therapeutic agents, e.g. toxins, to such cells
 PS Disclosure: Page 13; 31pp; English.
 CC The present sequence encodes a recombinant single chain peptide,
 CC 3G6-scrFv-streptavidin. The peptide is an antibody construct comprising
 CC the variable regions of the heavy and light chains of an antibody against
 CC disialoganglioside (GD2) as a single chain Fv fragment (scrFv). GD2 occurs
 CC in many tumours types including neuroblastoma, osteosarcoma and other
 CC soft tissue sarcomas, medulloblastomas, high grade astrocytomas,
 CC melanomas and small cell lung cancer. The peptide can be detectably
 CC labelled, preferably with 99mTc, for tissue imaging of cells expressing
 CC GD2. It can also be used to target delivery of a therapeutic or
 CC pre-therapeutic agent, such as a toxin, streptavidin (e.g. present
 CC sequence) or a pro-drug converting enzyme, to cells expressing GD2. The
 CC peptide may further comprise CD8 to facilitate the formation of
 CC GD2-targeted lymphocytes. T cells containing the peptide can also be used
 CC to target GD2-producing tumour cells.

SQ Sequence 1173 BP: 270 A: 334 C: 340 G: 229 T:
 Query Match 100.0%; Score 714; DB 1; Length 1173;
 Best Local Similarity 100.0%; Pred. No. 1.4e-187;
 Matches 714; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 agtattgtatgaccagactcccaattcctctgtatcagcaggagacagggttacc 60
 Db 1 AGTATTGTATGATACCCAGACTCCCAATTCTCTGTATATCAGCAGGAGACAGGGTTACC 60
 QY 61 ataacttcgaagccagtcagatgtagtaagtatggttggttaccacagaagcca 120
 Db 61 ATAACCTGCAAGGCCAGTCAGAGTGTAGTAATGATGTGGCTTGGTACCAACAGAAGCCA 120
 QY 121 gggcagttctccgaactgctgatactctgcacccaatcgctacactgagtcctgat 180
 Db 121 GGGCAGTCTCCGAACCTGCTGATATACTCTGCATCCAACTACACTGAGTCCCTGAT 180
 QY 181 cgttccactggcagtgatgggacggatttcacattccaccatcagcagctgtgcaggct 240
 Db 181 CGTTTCACTGGCAGTGGATATGGACGGATTTCACATTTCACCATCAGCAGCTGTGCAGGCT 240
 QY 241 gaagacctggcagtttattctctgcagcagattatagctcgtcggaggggggaccaag 300
 Db 241 GAAGACCTGGCAGTGTATTTCTGTACAGAGGATTATAGCTCGTGGAGGGGGGACCAAG 300
 QY 301 ctggaaataaagggtgagcgggttcaggcggagggtggtcgtggtggtgcgcatgcag 360
 Db 301 CTGGAATAAAGGTGGAGCGGTTTTCAGCGGAGGTGGCTCTGCGGGTGGCGGATGCGAG 360
 QY 361 gtcagggtgaagagtcagcagcctggcgtggtggcgcctccacagagcctgtccatcact 420
 Db 361 GTGAGGTGAAGCAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCCTGTCTCATCACT 420
 QY 421 tgcactgtctctgggttttcttaacaaattatggtgtacactgggttcgcaagcctcca 480
 Db 421 TGCACGTCTCTGGGTTTCTTATTAACCAATTATGGTGTACACTGGGTGGCGGATGCGAG 480
 QY 481 ggaagggtcgtgagtggtgggagtaataatggcgtggtggaagacacaaattataattcg 540
 Db 481 GGAAGGGTCTGGAGTGGCTGGGAGTAATATGGGTGGTGAAGCACAAATATATAATTCG 540
 QY 541 gctcttatgtccagactagcatcagcaagacaaactccagagaccagttttcttaaaa 600
 Db 541 GCTCTTATGTCCAGACTGAGCATCAGCAAGGACAACTCCAAGAGCCAAAGTTTCTTAAAA 600
 QY 601 atgaacagtctgaaactgatacacagccatgtactactgtccagtgcggggggttaac 660
 Db 601 ATGACAGCTGTGCAACCTGTATGACACACCCATGTACTGTGCCAGTCGGGGGGGTAAC 660
 QY 661 tacggctatgctttggactactgggttcaggaaacctcagtcaccgtctctcca 714
 Db 661 TACGGCTATGCTTTGGACTACTGGGTCAAGAACCTCAGTCAGCTCAGCTCTCTCA 714

RESULT 3
 T69218
 ID T69218 standard; DNA; 752 BP.
 AC T69218;
 DT 19-AUG-1997 (first entry)
 DE 2E12 sfv DNA.
 KW Modified sfv; cell adhesion; artificial ligand; tumour; cancer;
 KW gene therapy; retrovirus; vector; 2E12; ss.
 OS Not specified.
 PN WO9720048-A2.
 PD 05-JUN-1997.
 PR 27-NOV-1996; U19051.
 PR 30-NOV-1995; US-007755.
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Fell P, Hayden M, Ledbetter JA, Mittler R, Winberg G;
 DR WPI; 97-310604/28.
 PT Modified sfv molecule for mediating adhesion between cells -

PT contains antibody binding site and transmembrane domain of receptor,
 PT useful for enhancing immune responses to disease
 PS Example 1; Fig 12A; 69pp; English.
 CC A nucleic acid sequence (T69218) is provided of hybridoma 2E12
 CC (see also T69223-24). It is used in retrovirus vectors such as
 CC pLNC-2el2hlg1B7-1m (see also T69216) and pLNC-2el2g1GCD58GPI
 CC (see also T69217) that code for novel svf molecules which have been
 CC modified by connecting a transmembrane domain of a cell surface
 CC receptor to the antigen binding site of the molecule. This creates
 CC artificial ligands that stimulate adhesion between cells and
 CC enhance co-stimulatory activity during an immune response against
 CC disease. Such vectors can be used for gene therapy of e.g. cancer.
 CC Sequence 752 BP; 188 A; 182 C; 205 G; 177 T;
 SQ

Query Match 58.0%; Score 414.4; DB 1; Length 752;
 Best Local Similarity 76.9%; Pred. No. 3.7e-105;
 Matches 569; Conservative 0; Mismatches 141; Indels 30; Gaps 4;
 QY 4 attgtgatgacccagactcccaaatctctgtgtatcagcagagagaggtttaccata 63
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 7 ATTGTGTCACCAATCCAGCTCTTTGGCTGTCTGTAGTCAGAGAGCCACCATC 66
 QY 64 acctgaaggccagctgagtgtagtaagtgtgctt-----tggtaccaa 111
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 67 TCTGTGAGAGCCAGTGAAGTGTGAATATATATGTCACAAAGTTTAATGCGATGTCACAA 126
 QY 112 cagaagccagggcagctccgaaactctgatatactctgcatacctcaactgcactgga 171
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 127 CAGAAACAGGACAGACACCACTCTATCTCTGCTGCTACCAAGTAGATCTGGG 186
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 172 gtccctgatcgttctcactgagtgagtgatgggacgatttcaactttcacatcagcact 231
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 187 GTCCCTGCCAGTTTGTGCGAGTGGTCTGGGACAGACTTCAGCCCTCAACATCCATCCT 246
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 232 gtccagctgaagacccctggcagtttattctgtcagc-----aggattatagctcg 282
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 247 GTGGAGAGGATGATATGCAATATATTTCTGTGACAAAGTAGAAGGTTCCTTTGACG 306
 QY 283 ctggaggggggacccagctggaaataaaa--agtgagcgggttcagcagaggtggc 339
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 307 TTCGGTGGAGCACCAGCTGGAATCAACGAGGTGGCGGTGGCTCGGCGGTGGTGGG 366
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 340 tctggcgtggcgatcgaggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 399
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 367 TCGGCTGGCGCGGATCTCAGGTGAGTGAAGGAGTCAGGACCTGGCTGGTGGCGGCC 426
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 400 tcacagagcctgtccatcactgtcactgtctctgtgggttttcaataaactattgtgta 459
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 427 TCACAGAGCCTGTCCATCATGACCGCTCTCAGGGTTCATATTAACCGGCTATGTTGTA 486
 QY 460 cactgggttcgagcctccaggaagggtctgagtgagtgagtgagtgagtgagtgagtgag 519
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 487 AACTGGGTTGCCAGCCTCCAGGAAGGGTCTGAGTGGCTGGGAATGATATGGGTGAT 546
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 520 ggaagcaaatataattcggctcttatgtccagactgagcatcagcaaggacaaactcc 579
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 547 GGAAGCAGAGACTATAATTCAGCTCTCAATCCAGAGCTGAGCATCACCAAGGACAACTCC 606
 QY 580 aagaagcaagtttcttaaaataaagcagctgcaactgatgacacagcagcatgtactac 639
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 607 AAGAGCAAGTTTCTTTAAATAAAGAGTCTGCAAACTGATGACAGAGCCAGATATAC 666
 QY 640 tgtaccagctgggggtaactac-----ggctactcttgactactgggttcaagga 693
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 667 TGTGCCAGAGATGTTTATAGTAATTTTCAATTACTTATGTTATGGACTACTGGGGTCAAGGA 726
 QY 694 acctcagtcacccgtctcctc 713
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 727 ACCTCAGTCACCGTCTCCTC 746

RESULT 4
 Q51541
 ID Q51541 standard; DNA; 729 BP.
 AC Q51541;
 DE 23-MAY-1994 (first entry)
 DT Coding sequence of polypeptide with affinity for fluorescein.
 KW Monoclonal antibody; MAB; affinity; binding; antigen; diagnostics;
 KW therapy; imaging; purification; biosensors; ss.
 OS Synthetic.
 PN US5260203-A.
 PD 09-NOV-1993.
 PF 02-SEP-1986; 902971.
 PR 02-SEP-1986; US-902971.
 PR 19-JAN-1989; US-092110.
 PR 25-APR-1990; US-512910.
 PA (ENZO-) ENZO INC.
 PI Bird RE, Hardman K, Ladner RC;
 DR WPI; 93-367875/46.
 DR P-PSDB; R43680.
 PT Single chain poly:peptide for binding antigen - comprising light
 PT and heavy chain antigen binding portions linked by peptide linker
 PS Example 9; Figure 40; 78pp; English.
 CC The single chain polypeptide is derived from the mature light and
 CC heavy chains of a monoclonal antibody (MAB) and has affinity
 CC for a given antigen (fluorescein). It comprises a first
 CC polypeptide comprising the antigen binding portion of of the light
 CC chain variable region of an antibody and a second polypeptide
 CC comprising the antigen binding portion of the heavy chain variable
 CC region of an antibody and at least one peptide linker linking the
 CC first and second polypeptide chains. The resulting single chain
 CC polypeptide can be used in diagnostics, therapy
 CC (in vivo and in vitro), imaging, purifications and biosensors.
 CC This particular single chain binding molecule was designated
 CC 18-2-3/TRYS9 and contains one linker peptide.
 CC Sequence 729 BP; 177 A; 186 C; 178 G; 188 T;
 SQ

Query Match 57.2%; Score 408.4; DB 1; Length 729;
 Best Local Similarity 76.7%; Pred. No. 1.6e-103;
 Matches 557; Conservative 0; Mismatches 151; Indels 18; Gaps 4;
 QY 4 attgtgatgacccagactcccaaatctctgtgtatcagcagagagaggtttaccata 63
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 7 AATGTGTCACCCAGCTCCAGCAATCATGTCTGTCATCTCCAGGGGAAAGGTCCACATG 66
 QY 64 acctgaaggccagctcagagtgtagt---aatgatggctgtgtgtaccaaagagcca 120
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 67 ACCTGAGGCGCCAGCTCAAGTGAAGTTCCAGTTTCTGACTGGTACCAGCAAGATCA 126
 QY 121 gggcagctctccgaaactgtgtatatactctgcatacctcgcactgagtcacctgat 180
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 127 GGTGCTCCCCCAACTCTGGGTTTATGGCACATCACTTGGCTTCTGGAGTCCCTGCT 186
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 181 cgttcaactggcagtgatgagcaggtattcactttcacatcagcagcagcagcagcagc 240
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 187 CGGTTTCAGTGGCGGTGGTGGGAGCTCTTACTCTCTCAATCAGCAGTGGAGGCT 246
 QY 241 gaagactgacagttattctctcagcagagattatag-----ctcgctcagaggg 291
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 247 GAAGATGCTGCCACTTATTACTGCCAGCAGTACAGTGGTTACCCACTCAGTTCTGGTGGCT 306
 QY 292 gggaccaagctggaaataaagggtggaggggttcagggaggggtggtctctggtggtggt 351
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 307 GGGACCAAGCT---TAAAGAACTGCTGTTCTTCTTCTGAACAGCTGGCTCAGTTTCTGT 363
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 352 ggaatcagcaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtg 411
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 364 TCTCTGGATGTGAGCTGAAGGAGTCAAGACCTGTCTGTGGTGGCGCCCTCACAGAGCCTG 423
 QY 412 tcatcacttgactgtctctgggttttttaataaacaattatggtgtacatgggttgcg 471
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 424 TCCATCATTGCACTGTCTCTGGGTTTTCATTAAACCAACTATGTTGTACATGGGTTCGC 483

PT Immunoassay using single chain antigen binding mol. - as replacement
PT for labelled or immobilised antibody, are less immunogenic, easier
PT to engineer, more stable and less expensive
PS Example 9; Fig 40A-B; 78pp: English.

Computer-designed construct 18-2-3/TRV59 (W02192), encoded by the DNA sequence given in T36464, comprises the variable regions of the light and heavy chains of anti-fluorescein monoclonal antibody (Mab) 18-2-3 linked by a peptide designed to fit into a groove on the backside of the variable domain structure. The 18-2-3 VH and VL regions were obt'd. by PCR amplification of hybridoma RNA, and the DNA construct was inserted into vector pCX3703 and introduced into *E. coli*. 18-2-3/TRV59 was expressed as a single chain molecule. It exhibited biological binding activity equivalent in specificity and affinity to that of the original Mab.

Sequence	729 BP;	177 A;	186 C;	178 G;	188 T;
----------	---------	--------	--------	--------	--------

Query Match 57.2%; Score 408.4; DB 1; Length 729;
Best Local Similarity 76.7%; Pred. No. 1.6e-103;
Matches 557; Conservative 0; Mismatches 151; Indels 18; Gaps 4;

Qy	4	atgtgatgaccagagactcccaattctctgtgtatcagcaggagacagggttaccata	63
Db	7	AATGTGTCACCCAGTCCAGCAATCATCTGTCATCTCCAGGGAAAGGTCCACATG	66
Qy	64	acctgcaagccagtcagagtgtagt--aatgatgtggtctggtaccacaagaagcca	120
Db	67	ACCTGCAGGGCCAGCTCAAGTGTAAAGTTCAGTTACTTGCATGGTACCAGCAGAAGTCA	126

[illegible]

Qy	241	gaagaccctggcagttatttctctcagcaggattatag	-----ctcgctcggagg	291
Db	247	GAAGATGCTGCCACTATTACTGCCAGCAGTACAGTGGTTACCACTCAGTTTCGGTGT		306
Qy	292	ggaccaacgctgaaataaaaggtgaagcggtctcaggcggaggtggtctctgcgcgtg	gc	351

307	GGGACCAAGCT---TAAAGAAATCTGGTCTGTTCTTCTGTAACACAGCTGGCTCAGTTCTGCT	Db
352	ggatctgcaggtgcaggtgaaaggaggtcaggacctggcctggctggcgcgcctcacagagcctg	QY
364	TCCTGGATCTGCAGCTGAAGGAGTCAAGACCTTGCTGCTGGCGGCCCTCACAGAGCCTG	Db
412	tcctacacttgcactgctctctggagttttcattaaaccaattatggtatgatacactggggttcgc	QY

472	CGCCTCCAGGAAAGGCTGGAGTGGCTGGGAGTAAATATGGCTGGTGGAAACACAAAT	543	
484	CAGCTCCAGGAAAGGCTGGAGTGGCTGGGAGTAAATATGGCTGGTGGAAACACAAAT	543	
QY	472	cagcctccaggaaaggctggagtggtcgggagtaatatgctgctggtggaagcacaaat	531
Db	424	TCCATCACTTGCACCTGCTCTGGGTTTTCATTAAACCACTATGGTGTACACTGGGTTCCG	483
472	CGCCTCCAGGAAAGGCTGGAGTGGCTGGGAGTAAATATGGCTGGTGGAAACACAAAT	543	
484	CAGCTCCAGGAAAGGCTGGAGTGGCTGGGAGTAAATATGGCTGGTGGAAACACAAAT	543	

Qy	532	tataattcgctcttatgtccagacgagcatcagcaagagcaactcccaagagccaagt	591
Db	544	tataattcagctctcgtgccagctgagcatcagcaagagcaattcccaagagccaagt	603
Qy	592	ttcttaaaatgacacgtctgcaactgatgacacagccatgtactactgtgcca---	648
Db	604	ttcttataaaatgaacactctgcaattgatgacacagccattatctactgtgccaagca	663

50	QY	004	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	5
----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---

Dbb 724 TCCTAA 729

RESULT	7
Q05715	
ID	Q05715 standard; DNA: 729 BP.
AC	
CT	Q05715;
DT	07-JAN-1991 (first entry)
DE	18-2-3-TRY59.
KW	Monoclonal antibody; variable chain; anti-fluorescein; ss.
PN	US4946778-A.
PD	07-AUG-1990.

PF 19-JUN-1980; 299617.
 PF 02-SEP-1982; 902971.
 PR 08-SEP-1987; 05-092110.
 PR 19-JAN-1989; 05-299617.
 PA (GENE) GENEX CORP.617.
 PI Ledner, PC Bi-d RE, Hardman K;
 DR WPI: 90-260350/34.
 DR P-PSDB: F06483.

DK P-PSDB; R06483.
PT Single polypeptide chain binding molecules - having light chain
PT variable region of antibody linked by peptide to heavy chain
PT variable region.

PS variable region. 40; 68pp; English.
 CC The sequence was constructed from the variable regions of an anti-
 CC fluorescein MAb, 18-3-2, an IGM. The VL and VH cDNA sequences
 CC were synthesised by priming on RNA isolated from hybridoma cells.
 CC The sequence encodes a single chain binding molecule comprising
 CC the variable regions of heavy and light chains linked by peptides.
 CC See also Q05708-Q05719.
 SQ Sequence 725 BP; 177 A; 188 C; 176 G; 188 T;

SQ	Sequence	729 BP;	177 A;	188 C;	176 G;	188 T;
CC	acc	also	Q05700 Q05715.			

Query Match 57.0%; Score 406.8; DB 1; Length 729;
Best Local Similarity 76.6%;
Pred. NO. 4.5e-103;
Matches 556; Conservative 0; Mismatches 152; Indels 18

[illegible]

Db	67	ACCTGAGGCGCAGCTCAAGTGTAACTTCCAGTTTCACTTGCATCTGGTATCCACCAACAGTCA	126
Qy	121	gggcagctctccgaactctgtatatactctgcaatcgatcaactggagtcctcatg	180
Db	127	GGTGCCTCCCCCAACTCTGGGTTTATGCAACATCCAACTTGGCTTCTGGAGTCCCTGCT	186
Qy	181	cqgttcactcgagctgagatagagacggatttcactttccactcaagcactatgcaagct	240

QY	167	CGTT	CAGT	CGCG	GGTGGT	CTGGAG	CGCTCT	TACTCT	CACAT	CAGC	AGTGTG	GAGG	GT	246
Db	187	CGTT	CAGT	CGCG	GGTGGT	CTGGAG	CGCTCT	TACTCT	CACAT	CAGC	AGTGTG	GAGG	GT	246
QY	241	gaagac	ctgc	gcag	ttattt	ctctgc	gacg	gattat	ag-----	ctcgc	ctcg	gaggg	g	291
Db	247	GAAGAT	GCTGCC	ACTATT	ACTGCC	AGCAGT	ACAGTGGT	TACCC	ACTCAG	TTCCG	TGCT	GC	T	306

Qy	292	ggacccaagctgaaataataaaagtgagcgcgttcaggcgagggtgcctctggcgggtgac	351
Db	307	gggaccacagct---taagaatctggtttctgtttcttcaacagctgcctcagtttctgt	363
Qy	352	ggatcgcagctgcaggtggaaggagtcaggacctggcctggtggcgccctcaacagagcctg	411
Db	364	tcctctggaatctgcagctggaaggatcagacctgtcctggtggcgccctcacaagaccgt	423

[illegible]

Db 484 CAGCCICAGAAAGGTCGGAATGGCTGGGAGTAATATGGCTGGTGGAAACACAAT 543

Qy 532 tataatcggtcttatgtccagactgagcatcagcaaggacaactcccaagagccaagt 591

```

Db 544 TATAATTCAGCTCTCATGTCCAGATGAGCATCAACAAGACAATTCACAGGCGCAAGTT 603
Qy 592 tcttaaaatgaacagctctgaaactgatcacacagccatgtactactgtgcca---gt 648
Db 604 TCTTAAAAATGAACAGTCTGCAAAATGATGACACAGCCATATACTACTGTGCGCAACGA 663
Qy 649 cgggggggtaactcagctctgtcttctgactactcgggtcaaggaaacccctcagtcaccgtc 708
Db 664 CTGGAACCAATCTTTTACTATGCTATGACATATTGGGTCAAGGAACCTCAGTCACCGTC 723
Qy 709 tctca 714
Db 724 TCCTAA 729

RESULT 8
T69216
ID T69216 standard; DNA; 1528 BP.
AC T69216;
DT 19-AUG-1997 (first entry)
DE Retrovirus vector pLNC-2el2hlglb7-1tm encoding modified sfv.
KW Modified sfv; cell adhesion; artificial ligand; tumour; cancer;
KW gene therapy; retrovirus; vector; pLNC-2el2hlglb7-1tm; ss.
OS Synthetic.
PN W09720048-A2.
PD 05-JUN-1997.
PF 27-NOV-1996; U19051.
PR 30-NOV-1995; US-007755.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Fell P, Hayden M, Ledbetter JA, Mittler R, Winberg G;
DR WPI; 97-310604/28.
PT Modified sfv molecule for mediating adhesion between cells -
PT contains antibody binding site and transmembrane domain of receptor,
PT useful for enhancing immune responses to disease
PS Example 1; Fig 10A-B; 59pp; English.
CC A nucleic acid sequence (T69216) is provided of a modified sfv
CC encoded by retrovirus vector pLNC-2el2hlglb7-1tm. This comprises
CC a fusion between hybridoma 2E12 sfv (see also T69218), human IgG1
CC Fc portion (see also T69219) and the transmembrane domain from
CC human B7-1 (CD80) (see also T69220) in vector pLNC. This is an
CC example of novel sfv molecules that are modified by connecting a
CC transmembrane domain of a cell surface receptor to the antigen
CC binding site of the molecule. This creates artificial ligands that
CC stimulate adhesion between cells and enhance co-stimulatory
CC activity during an immune response against disease. The vectors
CC can be used for gene therapy of e.g. cancer.
SQ Sequence 1528 BP; 380 A; 434 C; 399 G; 315 T;

Query Match 56.3%; Score 401.8; DB 1; Length 1528;
Best Local Similarity 76.7%; Pred. No. 1.3e-101;
Matches 568; Conservative 0; Mismatches 142; Indels 31; Gaps 5;

Qy 4 attgtgatgaccagctcccaaatctctgtgtatcagcaggagacagggttaccata 63
Db 1 ATTGTGTCACCACTCAACAGTCTTCTTGGCTGTCTCTAGGTGACAGACCCACCATC 60
Qy 64 acctcagggccagctcagctgtgagtaagtgtggtc-----tggtaccacaa 111
Db 61 TCCTGCAGAGCCAGTGAAGTGTGAAATATATGTCACAAGTTAATGCAGTGGTACCAC 120
Qy 112 cag-aagcagggcagctctccgaaactcgtgatatactctgcataccatcgtacactgg 170
Db 121 CAGAAACCCAGGACAGACCCACCAACTCCTCATCTCTGCTGCATCCACAGTAGATCTGG 180
Qy 171 agtccctcagctcctcagtcggcagtcgggatgggacggatttcacacacagac 230
Db 181 GGTCCCTGCCAGGTTTGTAGTGGCAGTGGGTCTGGGACAGACTTCAGCCCTCAACATCC 240
Qy 231 tgtcagactgaagaccctggcagttattctgtcagc-----aggattatagctc 281
Db 241 TGTGAGGAGGATGATATTGCATGATATTCTGTCAGCAAAAGTAGGAAGGTTCTTTGGC 300

```

```

Qy 282 gctcggagggggggaccagctcgggaaataaaa---ggtggaggcggttcacgcgagggtgg 338
Db 301 GTTCGGTGGAGGACCAAGCTTGAATAATCAAAAGGGGTGGCGGTGGCTCGGCGGTGTGGTGG 360
Qy 339 ctcctggcgggtggcggatcgccaggtgcagggtgaaggagtcaggagcctcggcctgtggtggcc 398
Db 361 GTCGGGTGGCGGGGATCTCTCAGGTGCAGTGAAGAGTCAAGGAGCTGAGGACCTGGCTTGGTGGCGCC 420
Qy 399 ctacacagcctgtccatcactgtctctggtgttttcatttaaccaattatgtgt 458
Db 421 CTCACAGAGCCTGTCCATCATCATGACCGTCTCAGGGTCTCATTAACCGGCTATGTGTGT 480
Qy 459 acactgggttcggcagcctccagaaagggtcgtgagtggtggagtaatatgggtgg 518
Db 481 AAAGTGGGTTCGCCAGCCTCCAGAAAGGGTCTGAGTGGGTGGGAATGATATGGGTGA 540
Qy 519 tgggaagcacaaaattataattcggctcttattgtccagactgagcagcaggaagcaactc 578
Db 541 TGAAGCACAGACTATAATTAGCTCTCAAAATCCAGACTGAGCATCACCAAGGACAACTC 600
Qy 579 caagagccaaagtttcttaaaatgaacagctctgcaaaactgatcacacagccatgtacta 638
Db 601 CAAGAGCAAGTTTTCTTAAAAATGAACAGTCTGCARACTGATGACACAGCAGCATACTA 660
Qy 639 ctgtgcagctcggggggttaactac-----ggctatgcttggactactcgggtcgaagg 692
Db 661 CTGTGCCAGAGATGGTTATAGTAACCTTTTCATTACTATATGTTATGGACTACTGGGTCAAGG 720
Qy 693 aacctcagtcacgtctcctc 713
Db 721 AACCTCAGTCACCGTCTCCTC 741

RESULT 9
T69217
ID T69217 standard; DNA; 1510 BP.
AC T69217;
DT 19-AUG-1997 (first entry)
DE Retrovirus vector pLNC-2el2hlglCD58GPI encoding modified sfv.
KW Modified sfv; cell adhesion; artificial ligand; tumour; cancer;
KW gene therapy; retrovirus; vector; pLNC-2el2hlglCD58GPI; ss.
OS Synthetic.
PN W09720048-A2.
PD 05-JUN-1997.
PF 27-NOV-1996; U19051.
PR 30-NOV-1995; US-007755.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Fell P, Hayden M, Ledbetter JA, Mittler R, Winberg G;
DR WPI; 97-310604/28.
PT Modified sfv molecule for mediating adhesion between cells -
PT contains antibody binding site and transmembrane domain of receptor,
PT useful for enhancing immune responses to disease
PS Example 2; Fig 11A-B; 69pp; English.
CC A nucleic acid sequence (T69217) is provided of a modified sfv
CC encoded by retrovirus vector pLNC-2el2hlglCD58GPI. This comprises
CC a fusion between hybridoma 2E12 sfv (see also T69218), human IgG1
CC Fc portion (see also T69219) and the transmembrane domain from the
CC CD58 glyco-phosphoinositol (GPI) anchor (see also T69221) in vector
CC pLNC. This is an example of novel sfv molecules modified by
CC connecting a transmembrane domain of a cell surface receptor to the
CC antigen binding site of the molecule. This creates artificial
CC ligands that stimulate adhesion between cells and enhance co-
CC stimulatory activity during an immune response against disease.
CC Such vectors can be used for gene therapy of e.g. cancer.
SQ Sequence 1510 BP; 381 A; 422 C; 394 G; 313 T;

Query Match 56.3%; Score 401.8; DB 1; Length 1510;
Best Local Similarity 76.7%; Pred. No. 1.3e-101;
Matches 568; Conservative 0; Mismatches 142; Indels 31; Gaps 5;

Qy 4 attgtgatgaccagctcccaaatctctgtgtatcagcaggagacagggttaccata 63

```

```
Db 1 ATGTGCTACCCAAATCTCCAGTCTCTTTGGCTGTGTCTTAGTGCAGAGACCACCATC 60
QY 64 acctgcaaggccagtcagagtgagtaagtagtggtc-----tggtaccac 111
Db 61 TCCTGCAGAGCCAGTGAAGAGTGTGAATATATATGTCACAAAGTTAATGCAGTGGTACCAA 120
QY 112 cag-aagccaggcagtcctccgaaactgctgatatatactctgcatcgaatcgctacactgg 170
Db 121 CAGAAACCCAGAGCAGACGCCACCCAACTCCCTCATCTCTGCTGCATCCAACTAGAAATCGG 180
QY 171 agtccctgacgtctcactgagcagtgagatggagcagatttcactttccatcagcac 230
Db 181 GTGCCCTGCCAGGTTTGTAGTGGCAGTGGCTGGACAGACTTCAGCTCAACATCCATCC 240
QY 231 tgtcaggtggaagacctggcagctgtattctgtcagc-----aggtatagatc 281
Db 241 TGTGGAGGAGATGATATTCGAATGATTTCTGTCAGCAAGTAGGAAGGTTCTTTGGAC 300
QY 282 gctcggagggggacaaagctggaataaaa---ggtggaggcgggttcaggcggaggtgg 338
Db 301 GTTCGGTGGAGGACCAAGCTGGAATAATCAACGGGTGGCGGTGGCTCGGGCGGTGGTG 360
QY 339 ctctgaggtggcagtcgagctgaggtgaggaagtcaggaagtcagcctggcgtggcgc 398
Db 361 GTCGGGTGGCGGGAGTCTCAGGTGCAGCTGAAGGAGTCAGGACCTGGCCCTGGTGGCGCC 420
QY 399 ctacagagcctgtccatcactgacgtctgtctgtggttttcattacccaattatggtg 458
Db 421 CTCACAGAGCCTGTCCATCATGTCACCGTCTCAGGGTCTCTATTACCGGCTATGGTGT 480
QY 459 acactgggttcgcagcctccaggaagggtctggagtgctggagtgctggagtaatatgggtgg 518
Db 481 AAACCTGGGTTCGCAGCCTCCAGGAAGGCTCTGGAGTGGCTGGGAATGATATGGGGTGA 540
QY 519 tgaagcacaaataataatcgctcttattgtccagagctgagcagcagcagcagcagc 578
Db 541 TGAAGCACAGACTATAATTCAGCTCTCAATCCAGACTCAGACTCACCAGGACAACTC 600
QY 579 caagagccaaagtcttcaaaataaacagctctgcaactgagcagcagcagcagcagcagc 638
Db 601 CAAGGCCAAGTTTCTTAAATGAACAGCTCTGCAACTGATGACACAGCCAGATACTA 660
QY 639 ctgtgcagtcg9gggggtaactac-----ggctatgctttggactactggggtcaag 692
Db 661 CTGTGCAGAGATGGTTATAGTAACCTTCTATTACTATGTATGACTACTGTGGGTCAAGG 720
QY 693 aacctcagtcacgtctctcctc 713
Db 721 AACCTCAGTCACCGTCTCCTC 741
```

RESULT 10

```
T69222
ID T69222 standard; DNA; 824 BP.
AC T69222;
DE Murine anti-human CD28 hybridoma 9.3 sfv VL-VH gene fusion.
KW Modified sfv; cell adhesion; artificial ligand; tumour; cancer;
KW gene therapy; hybridoma 9.3; CD28; ss.
OS Mus sp.
FH Key
FT cds Location/Qualifiers
FT 7..819
FT /*tag= a
FT signal_peptide 7..66
FT /*tag= b
FT /note= "native 9.3 signal"
FT 68..816
FT /*tag= c
FT mat_peptide
FT /product= VL-VH fusion
FT 364..369
FT /*tag= d
FT /note= "Jk1"
```

```
FT mrna 781..819
FT /*tag= e
FT /note= "JH4"
PN W09720048-A2.
PD 05-JUN-1997.
PF 27-NOV-1996; U19051.
PR 30-NOV-1995; US-007755.
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
PI Fell P, Hayden M, Ledbetter JA, Mittler R, Winberg G;
DR WPI; 97-310604/28.
DR P-PSDB: W16688.
PT Modified sfv molecule for mediating adhesion between cells -
PT contains antibody binding site and transmembrane domain of receptor,
PT useful for enhancing immune responses to disease
PS Example 3; Fig 13; 69pp; English.
CC A nucleic acid sequence (T69212) codes for 9.3 sfv (W16688)
CC comprising a fusion between the light chain and heavy chain
CC variable regions of murine anti-human CD28 hybridoma 9.3. The
CC light chain V region is rearranged with a J gene homologous to
CC murine Jk2, and the heavy chain V gene is rearranged with a J gene
CC segment homologous to murine JH4. The construct was obt'd. by PCR
CC amplification (see also T69229-34) of the heavy and light chain
CC regions. Novel sfv molecules are modified by connecting a
CC transmembrane domain of a cell surface receptor to the antigen
CC binding site of the molecule. This creates artificial ligands that
CC stimulate adhesion between cells and enhance co-stimulatory
CC activity during an immune response against disease. Vectors
CC encoding the scfv can be used for the gene therapy of cancer.
SQ Sequence 824 BP; 191 A; 197 C; 232 G; 204 T;
```

Query Match 56.1%; Score 400.2; DB 1; Length 824;
Best Local Similarity 76.4%; Pred. No. 3.1e-101;
Matches 572; Conservative 0; Mismatches 136; Indels 39; Gaps 5;

```
QY 4 attgtgatgaccagactcccaaatctctgtatctatcagcaggagacaggttaccata 63
Db 70 ATTGTGCTCACCACCAATCTCCAGCTTCTTTGGCTGTGTCTCTAGGGCAGAGCCACCATC 129
QY 64 acctgcaaggccagtcagagtgagtaagtagtggtc-----tggtaccac 111
Db 130 TCCTGCAGAGCCAGTGAGAGTGTGTGAATATTATGTCACAAAGTTTATGCAGTGGTACCAG 189
QY 112 cagaagccaggcagtcctccgaaactgctgatatctatctgtcctccaatcgctacactgga 171
Db 190 CAGAAGCCAGGACAGCAGCCCAACTCCCTCACTTTTGTGTCATCCAACTGAGAACTTGGG 249
QY 172 gtcctgtatgcttactgagcagtgagtgatggagcaggtattcactttcaccatcagcact 231
Db 250 GTCCCTGCCAGGTTTGTAGTGGCAGTGGGTCTGGGACAAACTTCAGCCTCAACATCCATCCT 309
QY 232 gtgcaggtcgaagacactggcagttatttctgtcagc-----agattatagctcg 282
Db 310 GTGGACGAGATGATGTTGCAATGTATTCTGTACGAAAGTAGGAAGTTCCTTACAGC 369
QY 283 ctgcgaggggggaccacagctggaataaaa-----ggtggaggcgggttcacggcga 333
Db 370 TTCGAGAGGGGGACCAAGCTGGAATAAAACGGCTTCGGGTGGTGGCGGTTCTGGAGGT 429
QY 334 ggtggtctcggcgggtggcgatcg-----caggtgcaggtgaaaggagtcaggaacctggc 387
Db 430 GGCGGTTCAGCGGGCGGTGATCCCTCGGCTCAGGTGTCAGTGCAGCTGAAGGAGTCAAGGAGCTGCGC 489
QY 388 ctgggtggccctcacagacgtccatcactgtctcactgtctcgtgttttcattacc 447
Db 490 CTGGTGACGCCCTCACAGAGCCTGTCCATCACTTGTACTGTCTCTGGGTTTTCATTAGC 549
QY 448 aattatggtgtacactgggttcgccagcctccaggaagggtctgtgagtggtggagta 507
Db 550 GACTATGGTGTTCATTTGGGTTCGCCAGTCTCCAGGACAGGAGTGGAGTGGCTGGAGTA 609
QY 508 atatgggtcgtgggaagcacaaattataattcgggtctcttatgtccagactgagcatcagc 567
```

Db	610	ATATGGCGCTGGTGGAGCCAGCAATTAATATTCGGCTCTCATGTCACAGAAAGACGATCAGC	666
Qy	568	aagggacaactccaagagcccaagttttcttaaaaaatgaacagctctgcaaaactgatgacaca	627
Db	670	AAAGACAACCTCAAGGGCCCAAGTTTCTTAATAAATGAAGAGCTGCGAAGCTGATGACACA	729
Qy	528	gscatgtactactgtgccag---tcgggggggggtaatacagcgctatgctttggactactgg	684
Db	730	GCGGTGTATTACTGTGCCAGAGATAAGGGATACTCCTATTACTATTCTATGGACTACTGG	789
Qy	685	ggtcagaagaacctcaatcacctgctctc 713	
Db	790	GGTCAAGGAACCTCAGTCCCGTCTCCTC 818	
RESULT 11			
TL3739			
ID	TL3739	standard; DNA; 720 BP.	
AC	TL3739;		
DT	11-OCT-1996	(first entry)	
DE	Single chain binding molecule 18-2-3/TRY202', DNA.		
KW	Antibody engineering; single polypeptide chain binding molecule;		
KW	single chain antibody; SCA; heavy chain; light chain;		
KW	monoclonal antibody; MAB; immunoaffinity purification;		
KW	18-2-3/TRY202'; ss.		
OS	Chimeric Mus sp.		
OS	Chimeric synthetic.		
FT	Key	Location/Qualifiers	
FT	cds	619..621	
FT		/*tag= a	
FT		/transl_except= (619..621, aa:ile)	
FT	cds	658..660	
FT		/*tag= b	
FT		/transl_except= (658..660, aa:gly)	
PN	US5534621-A.		
PD	09-JUL-1996.		
PF	02-SEP-1986; 902971.		
PR	02-SEP-1986; US-902971.		
PR	02-SEP-1987; US-092110.		
PR	19-JAN-1989; US-299617.		
PR	25-APR-1990; US-512910.		
PR	01-APR-1993; US-040440.		
PR	06-JUN-1995; US-468992.		
PR	(ENZO-) ENZON LABS INC.		
PI	Bird RE, Hardman K, Ladner RC;		
PI	WPI: 96-333309/33.		
DR	P-PSDB; R99649.		
PT	Immunopurific. using single binding chain molecule including		
PT	antigen-binding parts of antibody light and heavy chain variable		
PT	regions connected by a linker - is smaller, stabler and less		
PT	expensive than complete antibodies		
PS	Example 9; Fig 39; 78pp; English.		
CC	A DNA construct (R99649) codes for single chain binding molecule		
CC	18-2-3/TRY202' (R99649), in which VL and VH regions of anti-		
CC	fluorescein monoclonal antibody 18-2-3 are joined by a peptide		
CC	linker composed primarily of alternating Gly and Ser residues, with		
CC	Glu and Lys residues to enhance solubility. The construct was		
CC	expressed in E.coli. Soluble, folded chain antibody (SCA) was		
CC	obtd. which was capable of exhibiting a biological binding		
CC	activity equivalent in specificity and affinity to that of a		
CC	monoclonal antibody. The SCA has the advantages of smaller size,		
CC	greater stability and reduced cost.		
SQ	Sequence 720 BP; 180 A; 175 G; 181 C;		

QY	64	acctgcaaggccagtcagagctgtgag- ---aatgatgtgcttgggtaccacaacagaagccca	120
DB	67	ACCTGCAAGGCCAGCTCAAGTGTAAAGTTCCAGTTACTTGCACCTGGTACCAGCAGAAAGTCA	126
QY	121	gggcagctctccgaactgctgatatactctgcatccaatcgctacaactggagtcacctgat	180
DB	127	GGTGCCTCCCCCAAACTCTGGGTTTATGGCACATCCAACTTGGCTTCTTGGAGTCCCTGCT	186
QY	181	cgtctcaactggcagtgatggacgagatttcactttccacatcagcactgtcgagct	240
DB	187	CGCTTCAGTCGGCGGTCTGGGACTCTTACTCTCTCACAATCAGCAGTGTGGAGGCT	246
QY	241	gaagacctggcagcttattctctcagcaggattatag- -----ctcgcctcgaggg	291
DB	247	GAAGATGCTGCCACTTATTACTCCAGCAGTACAGTGGTTACCCACTCAGTTCGGTGTCT	306
QY	292	ggaccaagctggaataataaagdtgagggcgttcaggcggttcaggcggtgctctggcgtgtgc	351
DB	307	GGGACCAAGCTTGAGCTGGAAGTAAA- -----TCTTCTGGTTCTGTTCGGAA	354
QY	352	ggatcgcagctgcaggtgaaggatcaggacctggcctgggtggcgccctccacagagcctg	411
DB	355	TCTAAATCTACTCAGCTGAAGGAGTCAGGACCTGTCTGTGGCGCCCTCACAGAGCCTG	414
QY	412	tccatcactgcactgctctcggtgttttcattaaaccaattatggttacactgggttcgc	471
DB	415	TCCATCACTTGCAGTCTCTCTGGGTTTTTCATTAACCACTATGTTGGTGTACACTGGGTTCCG	474
QY	472	cagcctccaggaagggctggagtggtcggtgggagtaatatgggtcggtggaagcacaaat	531
DB	475	CAGCTCCAGGAAAGGCTGGAGTGGCTGGGAGTAATATGGCTGTGTGGAAACACAAAT	534
QY	532	tataattcgcctcttatgtccagactgagcatcagcaagacaactcccaagagccaagt	591
DB	535	TATAAATCAGCTCTCATGTCCAGACTCAGCATCAGCAAAACAAATTCGAAGAGCCAAGTT	594
QY	592	ttcttaaaatgaacagctgcgaacctgatgacacagccatgtactactgtgcc- ---gt	648
DB	595	TTCTTAAANAATGAACAGCTCTGCAAAATGATGACACAGCCTATACTTACTGTGCCAACGA	654
QY	649	cggggggggttaactcacgctctgcttggactactggtgggtcaaggaaacctcagtcaccgtc	708
DB	655	CTGGAACGAATCTTTTACTATGTATGTATGGACTATTGGGGTCAAGAACCTCAGTCACCGTC	714
QY	709	tcctca 714	
DB	715	TCCTAA 720	
RESULT 12			
T36463			
ID	T36463 standard; DNA; 720 BP.		
AC	T36463;		
DT	13-NOV-1996 (first entry)		
DE	18-2-3/TRY202' single chain binding protein-encoding DNA.		
KW	Antibody engineering; monoclonal antibody; WAB;		
KW	Single chain antibody; immunoassay; fluorescein; 18-2-3/TRY202';		
KW	Single chain binding protein; ss.		
OS	Chimeric Mus musculus;		
OS	Chimeric synthetic.		
FT	Key	Location/Qualifiers	
FT	cds	558..660	
FT		/tag= a	
FT		/transl_except= (658..660, aa:Gly)	
PN	US5518899-A.		
PD	21-MAY-1996.		
PF	02-SEP-1986; 902971.		
PR	02-SEP-1986; US-902971.		
PR	02-SEP-1987; US-092110.		
PR	19-JAN-1989; US-299617.		
PR	25-APR-1990; US-512910.		
PR	01-APR-1993; US-040440.		
PR	06-JUN-1995; US-458988.		

QY	709	tcctca	714
Db	715	TCCTAA	720
RESULT	13		
T66143			
ID	T66143 standard:	cDNA:	737 BP.
AC	T66143;		
DT	14-JUL-1997	(first entry)	
DE	pUR.4124 insert encoding Vllys-linker-VHlys.		
KW	Bispecific antibody; bivalent antibody; double head;		
KW	antibody engineering; GOSA; glucose oxidase; Streptococcus sanguis;		
KD	ds,		
OS	Mus sp.		
FH	Key	Location/Qualifiers	
FT	cds	11..730	
FT		/tag= a	
FT		/product= Vllys-linker-VHlys	
FT	mat_peptide	11..334	
FT		/tag= b	
FT		/product= Vllys	
FT	misc_rna	335..379	
FT		/tag= c	
FT		/product= (Gly4Ser)3 linker	
FT	mat_peptide	380..727	
FT		/tag= d	
FT		/product= VHlys	
PN	W09714719-A1.		
PD	24-APR-1997:	E03605.	
PF	14-AUG-1996;	E03605.	
PR	16-OCT-1995;	EP-307332.	
PA	(UNIL) UNILEVER NV.		
PA	(UNIL) UNILEVER PLC.		
PI	DAVIS RJ, Van Der Logt CPE, Verhoeijen ME, Wilson S;		
DR	WPI; 97-245049/22.		
DR	P-PSDB; W09813.		
PT	Bispecific or bivalent antibody fragment analogues comprise complex		
PT	of two polypeptide chains containing respectively two heavy and two		
PT	light chain variable domains - useful therapeutically, in		
PT	immunoassays, for purification etc.		
PS	Example 1; Page 43-43; 121pp; English.		
CC	The EcoRI-HindIII insert (T66143) of plasmid pUR.4124 encodes a		
CC	polypeptide (W09813) in which a VL domain is linked to a VH domain		
CC	via a flexible linker. Novel DNA constructs (see also T66144-47)		
CC	code for the components (W09814-20) of bispecific or bivalent		
CC	antibodies. Such antibody fragment analogues comprise a binding		
CC	complex of two polypeptide chains, one with two VH domains in		
CC	series and the other with two VL domains in series, the two domains		
CC	being connected either directly or via a linker peptide, and the		
CC	complex contg. two pairs of variable domains (VH-A/VL-A and		
CC	VH-B/VL-B). The two variable domains may be different, producing		
CC	a bispecific construct or A and B are the same and the construct is		
CC	bivalent. A host cell is transformed with separate DNAs encoding		
CC	the polypeptides, optionally connected by linkers, and cultured		
CC	to produce both products, which are then allowed to combine to		
CC	form the antibody fragment analogues.		
SQ	Sequence	737 BP; 191 A; 200 G; 167 T;	
Query Match	55.8%;	Score 398.2;	DB 1; Length 737;
Best Local Similarity	75.2%;	Pred. No. 1.le-100;	
Matches	544;	Conservative	0; Mismatches 158; Indels 21; Gap:
QY	4	atgtgatgccaccagactcccaaatctcgttgttatcaggacaggagttaaccata	6
Db	14	ATCGAGCTCACCAGTCTCCAGCTCCCTTTCTGGGTCTGTGGAGAACTGTCAACATC	7
QY	64	acctgcgaagggccagtcagtgctgagtgtaatgatgctggtttgtaccacaagaagccagg	1
Db	74	ACATGTCGAGCAAGTGGAATTATTCACAATTAATTAGCATGTGTATCAGCAAAACAGGGA	1

```

QY 124 cagctcccaaacgtgtgatatactctgcatccaatcgctacactggagtcctctgtcgc 183
DB 134 AAATCTCCTCAGCTCTGCTGCTATTATACAACAACCTTAGCAGATGGTGTGCCATCAAGG 193
QY 184 ttcaactggcagtgatattggcgggattttcactttcaccatcagcactgtcaggtcgaa 243
DB 194 TTCAGTGGCAGTGGATCAGAGACACAATATTCTCAAGATCAACAGCTGCAACCTGAA 253
QY 244 gacctggcagttatttctgtcagcaggatt-----atagctcgcgtcggaggggg 294
DB 254 GATTTGGGAGTATTACTGTCAACATTTTGGAGTACTCCTCGGACGTTCGGTGGAGGG 313
QY 295 accaagctggaaataaaa---ggtgagcggttcaggcgagggtgtctctgtggtggtgc 351
DB 314 ACCAAGCTCGAGATCAACAGGGGTGGAGCGGTTTCAGGCGGAGGTGGCTCTGGCGGTGGC 373
QY 352 ggcctcaggtcagtggaagagtcaggacctggcctggcctggcgccctcacagagcctg 411
DB 374 GGATCGCAGGTGCAGCTGCAGGAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCCTG 433
QY 412 tccatcactgtcactgtctctgggtttttcattaaccaattatggtgtacactgggttcgc 471
DB 434 TCATCATCATGCACCTCTCAGGGTTCTCATTAACCGGCTATGGTGTAACTGGGTTCGC 493
QY 472 cagctcccgagaaagggtctggagtcggagtcggagtaatatggcgtggtggaagcacaat 531
DB 494 CAGCCTCCAGGAAAGGCTGGAGTGGCTGGGAATGATTGGGGTGTATGGAAACACAGAC 553
QY 532 tataattcggctcttatgtccagactgagcagcagcagcagcagcagcagcagcagcagc 591
DB 554 TATATTACAGTCTCAAAATCCAGACGTAGCATCAGCAAGCAACCTCCAGAGCAGGTT 613
QY 592 ttcttaaaatgacagctgtcgaactgtgacacagcagcagcagcagcagcagcagcagc 651
DB 614 TTCTTAAATGAACAGTCTGCACACTGATGATACACAGCCAGGTACTACTGTGCAGAGAG 673
QY 652 ggggttaactacggtatcttggactactcgtggggtcaggggtcagggacactcagcagc 711
DB 674 AGAGAT-----TATAGGCTTGACTACTGGGGCCAGGGACACCGTCAACCGTCTCC 724
QY 712 tca 714
DB 725 TCA 727

RESULT 14
T94606
ID T94606 standard; cDNA; 737 BP.
AC T94606;
DE EcorI-HindIII insert of pUR.4124 nucleotide sequence.
KW Multivalent antigen binding protein; variable domain; binding site;
OS diagnosis; therapy; medicine; immunoassay; target; pUR4124; ds.
FH Key
FT Location/Qualifiers
CDS
FT 11..730
FT /tag= a
FT /product= "VLLys-linker-VHlys"
FT mat_peptide
FT 11..334
FT /tag= b
FT /note= "VLLys"
FT misc_RNA
FT 335..379
FT /tag= c
FT /note= "linker"
FT mat_peptide
FT 380..727
FT /tag= d
FT /note= "VHlys"
FT
FT W09738102-A1.
PD 16-OCT-1997.
PF 26-MAR-1997; E01609.
PR 04-APR-1996; EP-302412.
PA (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.

```

```

PI Davis PJ, Van DER LOGT CPE, Verhoeven ME;
DR WPI: 97-512719/47.
PT P-PSDB; W35561.
PT Multivalent antigen binding protein - comprises antibody variable
PT domains, linked by their association, where each associated variable
PT domain pair forms an antigen binding site
PS Example 3; Pages 42-43; 100pp; English.
CC This is the nucleotide sequence of EcorI-HindIII insert of pUR4124. It
CC contains the DNA encoding VHLys-linker-VLLys sequences. This is used in
CC the construction of a novel multivalent antigen binding protein. This
CC novel protein comprises 2 polypeptides, each comprising, in series, 3 or
CC more variable domains of an antibody heavy or light chain respectively,
CC which are linked by association of the respective heavy and light chain
CC variable domains, where each associated variable domain pair forms an
CC antigen binding site. The multivalent antigen binding protein can be used
CC in medicine, diagnosis, therapy, an immunoassay method or for
CC purification. It may be used in targeting a tumour cell with natural
CC killer cells and cytotoxic agents. It may be used to target cell killing
CC enzymes, e.g. oxidases and peroxidases, to a species with which is an
CC antigenic component of a dental plaque, such as S. sanguis or S. mutans.
CC The antigen binding protein may conveniently be purified straight from
CC the supernatant using conventional purification techniques. As the
CC protein is self-assembling, there is no need to purify individual
CC subunits prior to coupling as in existing techniques.
SQ Sequence 737 BP; 191 A; 179 C; 200 G; 167 T;

```

```

Query Match 55.8%; Score 398.2; DB 1; Length 737;
Best Local Similarity 75.2%; Pred. No. 1.1e-100;
Matches 544; Conservative 0; Mismatches 158; Indels 21; Gaps 3;
QY 4 attgtgatccacagcagcctcccaaatctctgtgtatcagcagggagacaggggttaccata 63
DB 14 ATCGAGCTCACCCAGCTCTCCAGCCTCCCTTTCTGCTGCTGTGGGAGAACTGTCAACATC 73
QY 64 acctcgaagcagtcagagtggtgagtaaatgatgtggtgtgtgtgtgtgtgtgtgtgtgt 123
DB 74 ACATCTCGAGCAAGTGGGAATATTCCAAATATTATTAGCATGGTATCAGCAGAAACAGGGA 133
QY 124 cagctccgaaactcgtgatatactctgcatccaatcgctacactggagtcctctgtcgc 183
DB 134 AAATCTCCTCAGCTCTGCTGCTATTATACAACAACCTTAGCAGATGGTGTGCCATCAAGG 193
QY 184 ttcaactggcagtgatattggcgggattttcactttcaccatcagcactgtcaggtcgaa 243
DB 194 TTCAGTGGCAGTGGATCAGGAAACACAATATTCTCAAGATCAACAGCTGCAACCTGAA 253
QY 244 gacctggcagttatttctgtcagcaggatt-----atagctcgcgtcggaggggg 294
DB 254 GATTTGGGAGTATTACTGTCAACATTTTGGAGTACTCCTCGGACGTTCGGTGGAGGG 313
QY 295 accaagctggaaataaaa---ggtgagcggttcagcagcagcagcagcagcagcagcagc 351
DB 314 ACCAAGCTCGAGATCAACAGGGGTGGAGCGGTTTCAGGCGGAGGTGGCTCTGGCGGTGGC 373
QY 352 ggcctcaggtcagtggaagagtcaggacctggcctggcctggcgccctcacagagcctg 411
DB 374 GGATCGCAGGTGCAGCTGCAGGAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCCTG 433
QY 412 tccatcactgtcactgtctctgggttttcattaaccaattatggtgtacactgggttcgc 471
DB 434 TCATCATCATGCACCTCTCAGGGTTCTCATTAACCGGCTATGGTGTAACTGGGTTCGC 493
QY 472 cagctcccgagaaagggtctggagtcggagtcggagtaatatggcgtggtggaagcacaat 531
DB 494 CAGCCTCCAGGAAAGGCTGGAGTGGCTGGGAATGATTGGGGTGTATGGAAACACAGAC 553
QY 532 tataattcggctcttatgtccagactgagcagcagcagcagcagcagcagcagcagcagc 591
DB 554 TATATTACAGTCTCAAAATCCAGACGTAGCATCAGCAAGCAACCTCCAGAGCAGGTT 613
QY 592 ttcttaaaatgacagctgtcgaactgtgacacagcagcagcagcagcagcagcagcagc 651
DB 614 TTCTTAAATGAACAGTCTGCACACTGATGATACACAGCCAGGTACTACTGTGCAGAGAG

```

```

Db 614 TTCTTAAATGACACAGCTGTCACACTGATGACACACAGCCAGGTACTACTGTGCCAGAG 673
QY 652 ggggttaactacgctatgcttggactactgggttaaggaaacctcagtcacctctcc 711
Db 674 AGAGAT-----TATAGGCTTGACTACTTGGGGCCAAAGGACCAGGTCACCGTCTCC 724
QY 712 tca 714
Db 725 TCA 727

RESULT 15
Q51540
ID Q51540 standard; DNA; 720 BP.
AC Q51540;
DT 23-MAY-1994 (first entry)
DE Coding sequence of polypeptide with affinity for fluorescein.
KW Monoclonal antibody; MAb; affinity; binding; antigen; diagnostics;
KW therapy; imaging; purification; biosensors; ss.
OS Synthetic.
FH Key
FT misc_difference 619..621
FT /*tag= a
FT /transl_except= AAT encodes Isoleucine.
FT misc_difference 658..660
FT /*tag= b
FT /transl_except= GAA encodes Glycine.
PN US5260203-A.
PD 09-NOV-1993.
PF 02-SEP-1986: 902971.
PR 02-SEP-1986: US-902971.
PR 02-SEP-1987: US-092110.
PR 19-JAN-1989: US-299617.
PR 25-APR-1990: US-512910.
PA (ENZO-) ENZON INC.
PI Bird RE, Hardman K, Ladner RC;
DR F-PSDB; R43679.
PT Single chain poly:peptide for binding antigen - comprising light
PT and heavy chain antigen binding portions linked by peptide linker
PS Example 9; Figure 39; 78pp; English.
CC The single chain polypeptide is derived from the mature light and
CC heavy chains of a monoclonal antibody (MAb) and has affinity
CC for a given antigen (Fluorescein). It comprises a first
CC polypeptide comprising the antigen binding portion of of the light
CC chain variable region of an antibody and a second polypeptide
CC comprising the antigen binding portion of the heavy chain variable
CC region of an antibody and at least one peptide linker linking the
CC first and second polypeptide chains. The resulting single chain
CC polypeptide can be used in diagnostics, therapy
CC (in vivo and in vitro), imaging, purifications and biosensors.
CC This particular single chain binding molecule was designated
CC 18-2-3/TRY202' and contains one linker peptide.
SQ Sequence 720 BP; 180 A; 185 C; 175 G; 180 T;

```

```

Query Match 55.6%; Score 396.8; DB 1; Length 720;
Best Local Similarity 76.0%; Pred. No. 2.6e-100;
Matches 552; Conservative 0; Mismatches 147; Indels 27; Gaps 4;

```

```

QY 4 attgtatgacccagactcccaattcctgctgtatcagcagagacaggggtaccata 63
Db 7 AATGTGCTCACCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAGGTCCACATG 66
QY 64 acctcaagccagtcagactgtgagt---aatgatgtgctgtgtaccacagaagcca 120
Db 67 ACCTGCAGGGCCAGCTCAGAGTGAAGTTCAGTTACTTGACCTGGTACCAGCAGCAAGTCA 126
QY 121 gggcagctctccgaactctgtatatactctgcatccaatcgctacactggagtcctgat 180
Db 127 GGTGCTCTCCCAAACTCTGGGTTTATGGCACATCCAACTTGCTTCTGGAGTCCCTGCT 186
QY 181 cgttcactggcagtgatggacggatttcactttcaccatcagcactgtgcaggct 240

```

```

Db 187 CGCTTCAGTGGCGGTGGGTCTGGGACCTCTTACTCTCACAATCAGCAGTGTGGAGCT 246
QY 241 gaagacctggcagttatttctgtcagcagagattatag-----ctcgcctcgaggg 291
Db 247 GAAGATGCTGCCACATATTACTGCCAGCAGTACAGTGGTTACCCACTCAGCTTCGGTGT 306
QY 292 gggaccaagctggaaaaataaagggtggagcggttcagcgaagtggtctcgcggtggc 351
Db 307 GGGACCAAGCTTGAGCTGGAAGTAAA-----TCTTCTGTTCTTGGTTCCGAA 354
QY 352 ggatcgcgaggtgcaggtgaaggagtcaggaacctgcccctgggtggcgccctcacagagcctg 411
Db 355 TCTAAATCTACTCTAGCTGAAGGAGTCAGGACCTGCTCTGGTGGCCCTCAGAGAGCTG 414
QY 412 tccatcacttgcactgtctctgggttttcattaaaccaattatgggtgtacactgggttcgc 471
Db 415 TCCATCACTTGCACCGTCTCTGGGTTTTTCATTAAACCACTATGGTGTACACTGGGTTCGC 474
QY 472 cagcctccaggaaagggtctctgggtggagtggtggagtaatatgggtggtgggaagcacaat 531
Db 475 CAGCCTCCAGGAAAGGGTCTGGAGTGGCTGGAGTAATATGGCTGTGTGGAACACAAAAT 534
QY 532 tataattgggtcttattgtccagactgagcatcagcaggaagacaactccaagagccaagt 591
Db 535 TATAATTGAGCTCTCATGTCCAGACTGAGCATCAAGAACAAATTCACAGAGCCAAAGTT 594
QY 592 ttcttaaaatgaacagctctgcaactgatgacacagccatgtactactgtgcca---gt 648
Db 595 TTCTTAAAAATGAACAGTCTGCAAAATGATGACACAGCCATATATACTACTGTGCCAAACGA 654
QY 649 cgggggggtaactacggtctgcttggactactgggtcagggaacctcagtcacctc 708
Db 655 CTGGAACGAACTCTTTTACTATGCTATGGACTATTGGGGTCAAGGAACCTCAGTCACCGTC 714
QY 709 tcttca 714
Db 715 TCCTAA 720

```

Search completed: November 30, 1999, 14:11:38
Job time: 5223 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 14:00:09 ; Search time 383.27 Seconds
(without alignments)
4363.706 Million cell updates/sec

Title: US-08-940-544-4
Perfect score: 714
Sequence: 1 agtatgtgatgaccagac.....cctcagtcaccgtctctctca 714

Scoring table: IDENTITY_NUC

Searched: 3032314 seqs, 1171202697 residues

Database : EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*

54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	199.6	28.0	345	37	AA10291	AA10291 vt53a04.r
2	190.6	26.7	471	43	AI233978	AI233978 EST230666
3	152	21.3	328	31	AA295311	AA295311 EST100471
4	146.2	20.5	370	31	AA295093	AA295093 EST100400
5	141.6	19.8	345	32	AA335086	AA335086 EST39457
6	139.2	19.5	413	31	AA301347	AA301347 EST14279
7	139	19.5	438	39	AA880491	AA880491 vx40h05.r
8	138.6	19.4	393	20	T27593	T27593 EST100653.H
9	136.4	19.1	463	22	R69532	R69532 vj82d09.r1
10	136.4	19.1	352	31	AA301361	AA301361 EST14181
11	133.2	18.7	382	31	AA295786	AA295786 EST100987
12	130	18.2	349	32	AA379044	AA379044 EST91999
13	127.8	17.9	334	31	AA327254	AA327254 EST30547
14	126.8	17.8	357	32	AA361497	AA361497 EST71040
15	126.4	17.7	303	31	AA300788	AA300788 EST13648
16	125.8	17.6	516	22	R67559	R67559 yi42h11.r1
17	125	17.5	303	31	AA300891	AA300891 EST14031
18	123	17.2	381	20	T29114	T29114 EST69430.Hu
19	122.8	17.2	364	28	AA098196	AA098196 mm86h06.r
20	122.2	17.1	363	32	AA367405	AA367405 EST78511
21	122.2	17.1	432	35	AA544384	AA544384 vk33f04.r
22	121.4	17.0	517	63	AI989649	AI989649 ws26f12
23	120.8	16.9	823	29	AA170256	AA170256 ms87g10
24	119.8	16.8	364	20	T27579	T27579 EST100049
25	119.8	16.8	396	37	AA691311	AA691311 vs14f01.r
26	119.2	16.7	363	31	AA300651	AA300651 EST13764
27	119	16.7	301	32	AA377295	AA377295 EST9837
28	118.4	16.6	400	34	AA456778	AA456778 zw27f11.r
29	117.6	16.5	403	32	AA345486	AA345486 EST51505
30	117.4	16.4	395	20	T29112	T29112 EST69384.Hu
31	117	16.4	511	34	AA464794	AA464794 zx83h07.r
32	116.4	16.3	442	33	AA405415	AA405415 zu56d02.r
33	116.4	16.3	425	51	AI734035	AI734035 zu56d02.Y
34	115.8	16.2	332	20	T29342	T29342 EST77181.Hu
35	114.8	16.1	335	31	AA318377	AA318377 EST20620
36	114.8	16.1	405	32	AA360223	AA360223 EST69341
37	114.8	16.1	363	34	AA464313	AA464313 zx78c12.r
38	114.6	16.1	346	31	AA300582	AA300582 EST13427
39	113.4	15.9	282	31	AA300491	AA300491 EST13404
40	111.6	15.6	503	22	R28232	R28232 yhs7e08.r1
41	111.2	15.6	398	33	AA423447	AA423447 ve80a03.r
42	110.8	15.5	292	31	AA295508	AA295508 EST100751
43	110.2	15.4	418	34	AA515239	AA515239 ng69c07.s
44	109.6	15.4	469	49	AI645111	AI645111 ms87g10.Y
45	109.4	15.3	424	41	AI045896	AI045896 UI-R-CO-j

ALIGNMENTS


```

QY 124 cagtctccgaactgctgatatctctgcatactccaaatcgctacactggagctccctgatcgc 183
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 323 CAGTCTCTAAATGCTTTATTTCAAGACATCCAAACGACACAGGAGTCCCTGATCGC 264
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 184 ttcaactgcagtgatgagcagcagatttcacatttcacacatcagcagctgacagctgaa 243
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 263 TTCACAGGCAGTGGATCTGACACAGATTTTCACATTTCACCATCAGCAACATGTCAGGCTGAA 204
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 244 gacctggcagttatttctctgcagcagattatagct 280
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 203 GACCTGSCGTGTTATTACTGTATGTCAGTCTAACTCCT 167
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 3
AA295311 328 bp mRNA EST 18-APR-1997
LOCUS EST100471 Pancreas tumor I Homo sapiens cDNA 5' end similar to
DEFINITION immunoglobulin kappa light chain, VJ regions, mRNA sequence.
ACCESSION AA295311
NID 91947646
VERSION AA295311.1 GI:1947646
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 328)
AUTHORS Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,C.J., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.F., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
NATURE 377 (6547 Suppl), 3-174 (1995)
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT On Sep 1, 1995 this sequence version replaced.
Other_ESTs: THC169106
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavetigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the tigr Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. .328
/organism="Homo sapiens"
/db_xref="ATCC (inhost):190766"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"
/dev_stage="adult"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
EcORI; Site_2: XhoI"
73 a 93 c 78 g 77 t 7 others
BASE COUNT
ORIGIN

```

```

Query Match 21.3%; Score 152; DB 31; Length 328;
Best Local Similarity 70.6%; Pred. No. 9.8e-34;
Matches 197; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 4 attgtgatgacccagactcccaaatctcgtctgtatcagcagcagcagcaggtttaccacata 53
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 43 ATAGTGTGACGACAGNTTNCAGCCACCCCTGTCTGTCTTCAGGGGAAAGAGCCACCTCN 102
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 64 acctcaaggccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 123
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 103 TCTGCGAGGGCCAGTCAGAGTGTTAGCAGCACTTAGCTGTGTACCGAGAAACTGGC 162
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 124 cagtctccgaactgctgatatctctgcatactccaaatcgctacactggagctccctgatcgc 183
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 163 CAGGCTCCCGAGCTCCTCATCTATGTCATCATCCATCAGGNCACCTGGCATCCGCCAGG 222
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 184 ttcaactggcagtgatattggcagcagcagcagcagcagcagcagcagcagcagcagcagcag 243
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 223 TTCACTGGCAGTGGCTCTGGGACAGAGTTCACTCTCANCATCAGCATCTCGAGTCTGAA 282
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 244 gacctggcagttatttctctgcagcagcagcagcagcagcagcagcagcagcagcagcagcag 282
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 283 GATTTCAGTGTATTACTGTGTCAGCAGTATAATAACTGG 321
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

RESULT 4
AA295093 370 bp mRNA EST 18-APR-1997
LOCUS EST100400 Pancreas tumor I Homo sapiens cDNA 5' end similar to
DEFINITION similar to immunoglobulin kappa, variable region (GB:Y00640), mRNA
sequence.
ACCESSION AA295093
NID 91947582
VERSION AA295093.1 GI:1947582
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 370)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,C.J., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
NATURE 377 (6547 Suppl), 3-174 (1995)
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT On Sep 1, 1995 this sequence version replaced.
Other_ESTs: THC168243
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavetigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the tigr Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. .328
/organism="Homo sapiens"
/db_xref="ATCC (inhost):190766"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"
/dev_stage="adult"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
EcORI; Site_2: XhoI"
73 a 93 c 78 g 77 t 7 others
BASE COUNT
ORIGIN

```


DEFINITION	ACCESSION NID	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
------------	------------------	---------	----------	--------	----------	-----------	---------	-------	---------	---------

COMMENT 1

FEATURES
SOURCE

BASE COUNT	Query Match	336	9	396	9	454	178
ORIGIN	Best Local	58	1	118	1	178	178
	Matches						

QY	510 a
Dd	238 T

On Sep 12, 1996 this sequence version replaced gi:1404830.
Other ESTs: TNC167177
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Hum
an Index (<http://www.tigr.org/hgi/hgi.html>)

Query Match 19.1%; Score 136.4; DB 22; Length 463;
Best Local Similarity 70.3%; pred. No. 3.5e-29;

```

/db_xref="AACC (InnoSt):191/4/"
/db_xref="taxon:9606"
/clone_lib="Testis tumor"
/sex="male"
/dev_stage="adult"
/ret_vector="Vector: pLacZneo"

```

```

BASE COUNT
84 a 94 c 83 q 88 t 3 others
ECORI; Site_2: XhoI"
/note="Organ: testis; Vector: pBluescript SK-; Site_1:
/dev_stage= adult

```

BASE COUNT	84 d	94 C	83 g	88 E	3 Others
ORIGIN					

Query Match 19.1%; Score 136.4; DB 31; Length 352;
Best local similarity: 70.0%; Best No. 3; Best 20.

```
Matches 196; Conservative 0; Mismatches 83; Indels 1; Gaps 1;
```

— 1 —

64 acctgcaaggccagtcagagtgtagtaatgatgtggcttggtaccacagaagccaggg 123

.....

Dbb 183 AAAGCCCCTAAGCTCCTGATCTATGCTGCATCCAGTTTGCAAAGTGGGTCCCATCAAG 242

[illegible]

243 agacctggcagtttattttctatcagcaggaattatagctca 282

db 303 AGGTTTGCAAATTACTACTGTCAACAGNGTTACATACCG 342

RESULT 11
2205706

DEFINITION EST100987 Pancreas tumor I Homo sapiens cDNA 5' end similar to

ACCESSION AA295786

KEYWORDS EST.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

ADAMS, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,

White, O.; Sutton, G.; Blake, J. A.; Brandon, R. C.; Mah-Wal, C.; Clayton, R. A.; Cline, T. R.; Cotton, M. D.; Earle-Hughes, J.; Fine, J. D.

Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr.,

AA379044.1 GI:203126
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 349)
ADAMS, M.D., KERLAVAGE, A.R., FLEISCHMANN, R.D., FULDNER, R.A.,
BULT, C.J., LEE, N.H., KIRKNESS, E.F., WEINSTOCK, K.G., GOCAYNE, J.D.,
WHITE, O., SUTTON, G., BLAKE, J.A., BRANDON, R.C., MAN-WAI, C.,
CLAYTON, R.A., CLINE, T.R., COTTON, M.D., EARLE-HUGHES, J., FINE, L.D.,
FITZGERALD, L.M., FITZHUGH, W.M., FRITCHMAN, J.L., GEOGHAGEN, N.S.,
GLOCKE, A., GNEHM, C.L., HANNA, M.C., HEDBLUM, E., HINKLE, P.S., JR.,
KELLEY, J.M., KELLEY, J.C., LIU, L.-I., MARMARO, S.M., MERRICK, J.M.,
MORENO-PALANCA, R.F., McDONALD, L.A., NGUYEN, D.T., PELLIGRINO, S.M.,
PHILLIPS, C.A., RYDER, S.E., SCOTT, J.L., SAUDEK, D.M., SHIRLEY, R.,

TITLE
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence

JOURNAL
Nature 377 (6547 Suppl.), 3-174 (1995)

MEDLINE
96026280

On Sep 12, 1996 this sequence version replaced gi:1393183.
Other_ESRs: IMC169106
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavetigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene

```

seq primer: M13 reverse.
Location/Qualifiers
    1..349
    /organism="Homo sapiens"
FEATURES
    source

```

```

/obj_xref=Acc (innocst);1633/3
/obj_xref="taxon:9606"
/clone_lib="skin tumor I"
/dev_stage="adult"
/note="Organ: skin; Vector: pBluescript SK-; Site_1:
      ECORI; Site_2: XhoI"
      84 a 101 c 86 g 77 t 1 others
BASE COUNT
ORIGIN

Query Match 18.2%; Score 130; DB 32; Length 349;
Best Local Similarity 70.7%; Pred. NO. 2.2e-27;
Matches 188; Conservative 0; Mismatches 75; Indels 3; Gaps 1;

QY 4 attgtgatccagactcccaattcctctgtatcacagagagacagggttaccata 63
|||||.....||||| ||||| ||||| ||||| ||||| |||||
Db 19 ATTGTGTGACTCAGTCTCCAGCACCTCTGTTGTCTCAGGGGAAAGACGCCCTC 78
|||||.....||||| ||||| ||||| ||||| ||||| |||||
QY 64 acctgcaaggccagtccagagtgtgagtaatg---atgtggcttgggtaccacaagaacca 120
|||||.....||||| ||||| ||||| ||||| ||||| |||||

```

121 gggcagctctccgaaactgtgatatactctgcattccaatcgctacactgagtcacctgat 180
122 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
139 GGCAGGCTCCCAAGGTCTTCATCTATGGTGATCCAGCAGGCCATGGCATCCCCAGAC 198
140 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 cgcttcaactcgacagtggatatggacgggatttcactttcacctcagcaactgtgcaggct 240
182 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 AGTTCAGTGGCAGTGGGTCTGGACAGACTTCACCTCTACCATTACAGCACTCGAGSCT 258
200 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13

AA327254	AA327254	334 bp	mRNA	EST	20-APR-1997
LOCUS	EST30547	Colon I Homo sapiens	cdNA 5' end similar to immunoglobulin kappa light chain, V region (GB:L01279), mRNA sequence.		
DEFINITION	AA327254				
ACCESSION	AA327254.1	GI:1979498			
NID	91979498				
VERSION	AA327254.1	GI:1979498			
KEYWORDS	EST.				
SOURCE	human.				

ORGANISM

REFERENCE
1 (bases 1 to 334)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Feldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, A.R., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pellingrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.B., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE

JOURNAL
MEDLINE
96026280
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
On Sep 12, 1996 this sequence version replaced g1:1326576.

Contact: Kerlavade. AR

Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel.: 3018699056
Fax: 3018699423

Email: arkerlav@tiar.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES

```

1. 334
/organism="Homo sapiens"
/db_xref="rACC (inhost):127923"
/db_xref="taxon:9606"
/clone_lib="Colon I"
/dev_stage="adult"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI"
78 a 88 c 83 g 84 t 1 others
BASE COUNT

```

Query Match

Query Match 17.9%; Score 127.8; DB 31; Length 334;
Best Local Similarity 70.1%; Pred. No. 9.6e-27;
Matches 185; Conservative 0; Mismatches 78; Indels 1; Gaps 1

Qy 4 attgtatgacccagactcccaattcctgcttctatctatctagcagagacagqqtaccata 63

Db	71	ATCCAGATGACGAGCTNCCATCCTCCCTGGTCTGCACTGTAGGACACAGAGTTACCATC	130
Qy	64	acctcgaaagcgaagtcagagtgtagtaatatgatgtggctgtgtaccacaagacgacgag	123
Db	131	ACCTGCGCGGCAAGTCAGAGTGTAGAGGTATTATAATTTGGTATCAGCAGAAACACAGG	190
Qy	124	cagctctccgaacctgctgatatactatcgatccaatcgctacactgaggtccctgatcgc	183
Db	191	AAAGCCCTTAAGTCTCTGATCTATGCTGGTCCAGTTTGCAAAGTGGGTCCCATCAAG	250
Qy	184	ttcaactggcagtgagatattgggacgagatttcactttcacatcagcactgtgaagcc-tga	242
Db	251	TTCAGTGGCGGTGGATTGTGAGACAGATTTTCAGTCTCACCATCAACAGTCTGCAATCTTGA	310
Qy	243	agacctggngagttattctgtca	266
Db	311	AGATTCTGCACTTACTACTGTCA	334

RESULT 14

AA361497	AA361497	357 bp	mrna	EST	21-APR-1997
LOCUS	EST17040	T-cell lymphoma Homo sapiens	cdna 5'	end similar to	
DEFINITION	immunoglobulin kappa light chain, V region, mRNA sequence.				
ACCESSION	AA361497				
NID	G2014052				
VERSION	AA361497.1	GI:2014052			
KEYWORDS	EST.				
SOURCE	human.				

SOURCE
ORGANISM
Homo sapiens
Italian.

ORGANISM HOMO SUPRENS Chordata; Craniata; Vertebrata; Mammalia; Eukaryota; Metazoa; Eutheria; Primates; Catarrhini; Hominidae; Homo

REFERENCE 1 (bases 1 to 357)

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.E.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Frithman, J.L., Geoghegan, N.S., Glöck, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S.Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Werrick, J.M., Moreno-Palancques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D., Kunsch, C., HungJun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

7777.E

JOURNAL
Nature 377 (6547 Suppl.), 3-174 (1995)
based upon 83 million nucleotides of cDNA sequence
initial assessment of human gene diversity and expression patterns

MEDLINE 96026280

COMMENT On Sep 12, 1996 this sequence replaced qi:1404716.

Other ESTs: THC167177

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seg primer: M13 Reverse.

FEATURES

source
J. .357

```
/organism="Homo sapiens"
```

```
/db_xref="ATCC (inhost):165882"
```

```
/clone_lib="T-cell lymphoma"
/cell_type="T-lymphocyte"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 88 a 96 c 84 g 87 t 2 others
ORIGIN
Query Match 17.8%; Score 126.8; DB 32; Length 357;
Best Local Similarity 70.7%; Pred. NO. 1.9e-26;
Matches 195; Conservative 0; Mismatches 79; Indels 2; Gaps 2;
QY 4 attgtgagaccagactcccaaatctctgtgtatcagcagagacagggttaccata 63
DB 69 ATCCAGATGACCCAGTCTCCATCCCTCTGTGTGATCTAGGAGACAGATCACCATC 128
QY 64 acctcaagggccagtcagagtgatgagtgatggtgtgtaccacagaagccagg 123
DB 129 ACTTCCCGGGAAGNAGAGAAATTAACAACCTATTAAATTGGTATCAACAGAACCCAGG 188
QY 124 cagctccgaactgctgatatctctgcacccaatc-gctacacgtgagtcctgtatcg 182
DB 189 AAGGCCCTAAGCTCTCATCTACGGTCATCCAGTCTGCAAGTGGGGTCCCATCAAG 248
QY 183 ctctcagcagtgatgagtgagga-gga-ttccactttccaccatcagcactgtcaggctg 241
DB 249 GTTCAGTGGCAGTGATCTGGGACAGATTTTCACTCTCANCATCAGCGGTCTGCAACCTG 308
QY 242 aagacctgagctgtattcttctgtcagcaggattata 277
DB 309 AAGATTTTGAACCTTACTACTGTCAACAGAGTTACA 344
RESULT 15
AA300788 303 bp mRNA EST 18-APR-1997
LOCUS AA300788
DEFINITION EST13648 Testis tumor Homo sapiens cDNA 5' end similar to
immunoglobulin kappa light chain, V region, mRNA sequence.
ACCESSION AA300788
NID g1953120
VERSION AA300788.1 GI:1953120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 303)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palaoques,R.F., McDonald,D.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397991.
Other ESTs: THC167177
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
```

9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES

source
1..303
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="AYC (inhost):192224"
/db_xref="taxon:9606"
/map="5q32-33"
/clone_lib="Testis tumor"
/sex="male"
/dev_stage="adult"
/note="Organ: testis; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

BASE COUNT 69 a 88 c 72 g 72 t 2 others
ORIGIN

Query Match 17.7%; Score 126.4; DB 31; Length 303;
Best Local Similarity 71.2%; Pred. No. 2.3e-26;
Matches 166; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 4 attgtgagaccagactcccaaatctctgtgtatcagcagagacagggttaccata 63
DB 71 ATCCAGATGACCCAGTCTCCATCCCTCTGTGTGATCTAGGAGACAGATCACCATC 130
QY 64 acctgcaagggccagtcagagtgatgagtgatggtgtgtaccacagaagccagg 123
DB 131 ACTTCCCGGGAAGTCAACAGTATAGCACCCTATTAAATTGGTATCAGCAGAACCCAGG 190
QY 124 cagctccgaactgctgatatctctgcacccaatc-gctacacgtgagtcctgtatcg 183
DB 191 AAGGCCCTAAGCTCTCATCTACGGTCATCCAGTCTGCAAGTGGGGTCCCATCAAG 250
QY 184 ctctcagcagtgatgagtgagga-gga-ttccactttccaccatcagcactgtgca 236
DB 251 TTCAAGTGTCAAGTGTGGGACAGATTTTCACTCTCANCATCAGCGGTCTGCA 303

Search completed: November 30, 1999, 14:00:11
Job time: 4539 sec

QY 64 acctcgaaggccagtcacagtgatgagtaagtgtggtggttaccacaagaagccagg 123
Db 67 ACCTCAAGGCCAGTCAGAGTGTGAGTACTGATGTAAGTTGGTACCAACAGAACCCAGG 126
QY 124 cagtcctcgaagctgctgatatactctcctccatccatcctacactgagtcctgctgacgc 183
Db 127 CAGTCTCTAACTACTGATATACATGATCAATCCATCCCTACAGTGGAGTCCCTGATCCG 186
QY 184 ttcactggcagtggtgatggagcagcagatttcactctccaccatcagcactgtgcagtgctaa 243
Db 187 TTCAGTGGCAGTGGATATGGGAGCGATTTCATCTTCACCATCAGCACGTGTCAGGCTGAA 246
QY 244 gacctggcagttatttctgtcagcagagattatag-----ctcgtcggagggggg 294
Db 247 GACCTGGCAGGATTATTTCTGTGACGAGGATTATAGGCTCCATTCCAGCTTCGGCTCGGG 306
QY 295 accaagctggaataaaagggtggaagcgggtcagcgagatgctctggcgggtggcgga 354
Db 307 ACAGAGTTGGAAATAAA---GTCACAAATCCTCAGGATCTGGCTCCGAAATCCAAAGC 363
QY 355 tcgcaggtgcaggtgaaaggagtcaggacctgcgctgggtggcgccctcacagagcctgtcc 414
Db 364 ACGCAGGTCAAACTCGAGGAGTCTGGCCTGGGATATTGCAGCCCTCCAGACCCCTCAGT 423
QY 415 atcaattgcactgtctctgggttttcatttaaccaattat-----ggtgtacactgggtt 468
Db 424 CTGACTTGTCTTCTCTGGGTTTTCACGTGAGCACCTCTGTGATGGGTGAGCTGGATT 483
QY 469 ggcagcctccaggaaggggtctgagtggtggagtaatatggcgtggtgggaagcaca 528
Db 484 CGTCAGCCTTCAGGAAGGGTCTGAGTGGCTGGGACACATTTATGGGATGATGACAAA 543
QY 529 aattataattggctcttatgttccagactgagcatcagcaggaagcaactccaagagccaa 588
Db 544 CACTATAACCCATCCTCAAGAGCGGCTCAATCTCCAAGGATACCTCCACCAACACAG 603
QY 589 gttttcttaaaatgaacagctcgaagctgacacagcagcactgtactactgtgccaagt 648
Db 604 GTATTCTCAAGATCACAGTGTGGACACTGCAGATATGCGCATATCTTCTGTGCTCGA 663
QY 649 cg-----ggggggttaactacgctctgttgggactcgtgggtcgaaggaaacctca 699
Db 664 AGAGTCTCTATGTAATGGGGGACTATGCTATGGGACTACTGGGGTCAAGGAACCTCA 723
QY 700 gtcacgtctcctca 714
Db 724 GTCACCGTCTCCTCA 738

RESULT 2

US-08-752-844-36
; Sequence 36, Application US/08752844
; Patent No. 5935821
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; APPLICANT: Foon, Kenneth A.
; APPLICANT: Chatterjee, Sunil K.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE
; TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/752,844
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 30414-20002.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-752-844-36

Query Match 41.6%; Score 297; DB 4; Length 351;
Best Local Similarity 91.8%; Pred. No. 2.7e-80;

Matches 325; Conservative 0; Mismatches 26; Indels 3; Caps 1;

QY 358 caggtgcagtggaaggagtcaggacctggcctggtggcgccctcacagagcctgtccatc 417
Db 1 CAGGTNCAGCTGAAGGAGTCAGGACCTGGCTGGTGGCACCCCTCACAGAGCCCTGTCCATC 60
QY 418 acttgcactgtctctgggttttcatttaaccaattatggtgtacacgggttgcagcct 477
Db 61 ACATGCACTGTCTCTGGGTCTCATTTATCAGATATAGTGTACACTGGGTTCGCCAGCCT 120
QY 478 ccaggaaagggtctggagtggtggtggagtaatatggcgtggtggagcaacaataataat 537
Db 121 CCAGGAAGGGTCTTGAGTGGCTGGGAATGATATGGGGTGGTGGAAACACAGACTATAAT 180
QY 538 tcggctcttatgtccagactgagcatcagcaaggacaactccaagagcccaagtttttta 597
Db 181 TCAGCTCTCAATCCAGACTGAGCATCAGCAAGGACAACTCAAGAGCCCAAGTTTCTTA 240
QY 598 aaaatgaacagctgcacaaactgatgacacagcagcactgtactactgtgcagtcgggggggt 657
Db 241 AAAATGAACAGTCTGCAAACTGATGACACAGCCATGACTACTGTGCTGCA---GAGATGGT 297
QY 658 aactacgctatgcttggactactgggtcgaaggaaacctcagtcacccgtctcc 711
Db 298 TACTACACTATGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCC 351

RESULT 3

US-08-190-199A-60
; Sequence 60, Application US/08190199A
; Patent No. 5830663
; GENERAL INFORMATION:
; APPLICANT: EMBLETON, Michael J.
; APPLICANT: GOROCIOV, Guy
; APPLICANT: JONES, Peter T.
; APPLICANT: WINTER, Gregory P.
; TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,199A
FILING DATE: 13-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/01483
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9212419.7
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9117352.6
FILING DATE: 10-AUG-1991
SEQUENCE CHARACTERISTICS:
LENGTH: 708 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..705
US-08-190-199A-60

Query Match 39.8%; Score 284; DB 3; Length 708;
Best Local Similarity 95.1%; Pred. No. 3e-76;
Matches 293; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 358 caggtgcaggtgaaggagtcagagaccctggcctgtgctgcgcctccacagagcctgtccatc 417
Db 1 CAGGTGCAGCTGAAGGAGTCAAGGACCTGGCTGTGGCGCCCTCACAGAGCCTGTCCATC 60
QY 418 acttgactctctgtggttttcaattacaaattatgtgtacactgggttcgcagcct 477
Db 61 ACTTGCACTCTCTGTGGTTTTCATTACCACTATGGTGACACTGGGTTCGCCAGCCT 120
QY 478 ccaggaagggtctggagtgctgggagtaataatggctgggaggaagcacaataataat 537
Db 121 CCAGGAAGGCTCTGGAGTGGCTGGGAGTATATGGCTGGTGAAGCACAATAATAAT 180
QY 538 tcggctcttatgtccagactgagcatcagcaagagcaactccaagagccaaagtcttcta 597
Db 181 TCGGCTCTCATGTCCAGACTGAGCATCAGCAAGAGCAACTCCAAGAGCAAGTTTCTTA 240
QY 598 aaatgaacagctctgcaaaactgatgacacagccatgtactactgtgccagtcgggggggt 657
Db 241 AAAATGAACAGTCTGCAAACTGATGACACAGCCCATGTACTGTGCGCAGAGATCGGGGG 300
QY 658 aactacgg 665
Db 301 GCTTACTG 308

RESULT 4
US-08-190-199A-66
Sequence 66, Application US/08190199A
Patent No. 5830563
GENERAL INFORMATION:
APPLICANT: EMBLETON, Michael J.
APPLICANT: GORCHOV, Guy
APPLICANT: JONES, Peter T.
APPLICANT: WINTER, Gregory P.
TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,199A
FILING DATE: 13-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/01483
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9212419.7
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9117352.6
FILING DATE: 10-AUG-1991
SEQUENCE CHARACTERISTICS:
LENGTH: 669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..666
US-08-190-199A-66

Query Match 39.8%; Score 284; DB 3; Length 669;
Best Local Similarity 95.1%; Pred. No. 2.9e-76;
Matches 293; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 358 caggtgcaggtgaaggagtcagagaccctggcctgtgctgcgcctccacagagcctgtccatc 417
Db 1 CAGGTGCAGCTGAAGGAGTCAAGGACCTGGCTGTGGCGCCCTCACAGAGCCTGTCCATC 60
QY 418 acttgactctctgtggttttcaattacaaattatgtgtacactgggttcgcagcct 477
Db 61 ACTTGCACTCTCTGTGGTTTTCATTACCACTATGGTGACACTGGGTTCGCCAGCCT 120
QY 478 ccaggaagggtctggagtgctgggagtaataatggctgggaggaagcacaataataat 537
Db 121 CCAGGAAGGCTCTGGAGTGGCTGGGAGTATATGGCTGGTGAAGCACAATAATAAT 180
QY 538 tcggctcttatgtccagactgagcatcagcaagagcaactccaagagccaaagtcttcta 597
Db 181 TCGGCTCTCATGTCCAGACTGAGCATCAGCAAGAGCAACTCCAAGAGCAAGTTTCTTA 240
QY 598 aaatgaacagctctgcaaaactgatgacacagccatgtactactgtgccagtcgggggggt 657
Db 241 AAAATGAACAGTCTGCAAACTGATGACACAGCCCATGTACTGTGCGCAGAGATCGGGGG 300
QY 658 aactacgg 665
Db 301 GCTTACTG 308

RESULT 5
US-08-634-783A-3
Sequence 3, Application US/08634783A
Patent No. 5861276
GENERAL INFORMATION:
APPLICANT: KWAK, JU-WON
APPLICANT: HAN, MOON-HI
APPLICANT: CHOI, BYUNG-KWON
TITLE OF INVENTION: c-DNAS ENCODING MURINE ANTIBODY

;; TITLE OF INVENTION: AGAINST HUMAN PLASMA APOLIPOPROTEIN B-100
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: DILWORTH & BARRESE
;; STREET: 333 EARLE OVINGTON BLVD.
;; CITY: UNIONDALE
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 11553
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/634,783A
;; FILING DATE: 19-APR-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BARRESE, ROCCO S.
;; REGISTRATION NUMBER: 25,253
;; REFERENCE/DOCKET NUMBER: 685-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 516-228-8484
;; TELEFAX: 516-228-8516
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 693 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
US-08-634-783A-3

Query Match 39.8%; Score 284; DB 3; Length 693;
Best Local Similarity 87.4%; Pred. No. 2.9e-76;
Matches 311; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Oy 359 aggtgcaggtgaaggagtcaggacccctggctggcgccctcacagagccctgtccatca 418
Db 2 AGGTGCAGCTGGTGGAGTCAGACACCTGGCTGGTGGCGCCCTCACAGAGCCCTGTCCATCA 61
Oy 419 cttgcactgtctctgggttttcattaaacaaattatggtgtacactgggttcgccagcctc 478
Db 62 CBTGCACCTGCTCAGGGTTCTCTATTAAACCGACTATGTTGTAAGCTGGATTGCCAGCCTC 121
Oy 479 caggaagggtctggagtggctgggagtaattatggctggtggaagcacaattataatt 538
Db 122 CAGGAAGGGTCTGGAGTGGCTGGGAGTAATTGGGCTGGTGGAGCACATCTCTATAATT 181
Oy 539 cggctcttatgtccagactgagcatcagcaaggacaactccaagagcccaagtttctctaa 598
Db 182 CAGCTCTCAAGTCCAGACTGAGCATCAACAGGACAACTCCAAGAGCCCAAGTTTCTTAA 241
Oy 599 aaatgaacagtcgtcaaaatgatgacacagccatgtactactgtgccagtcgggggta 658
Db 242 AAATGAACAGTCTGCACACTGTATGACACAGCCATGTACTGTCTCAAAACATGAGGATA 301
Oy 659 actacggtctgtcttgactactgggtcaggaacactcagtcacccgtctcctca 714
Db 302 GGTACGACTGGTACTTCGATGTCTGGGGCGCAGGGACACCGGTCCCGTCTCTCTCA 357

RESULT 6
US-08-752-844-32
; Sequence 32, Application US/08752844
; Patent No. 5935821
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; APPLICANT: Foon, Kenneth A.
; APPLICANT: Chatterjee, Sunil K.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE

;; TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
;; NUMBER OF SEQUENCES: 66
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: PALO ALTO
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/752,844
;; FILING DATE:
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schiff, J. Michael
;; REGISTRATION NUMBER: 40,253
;; REFERENCE/DOCKET NUMBER: 30414-20002.21
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 813-5600
;; TELEFAX: (415) 494-0792
;; TELE: 706141
;; INFORMATION FOR SEQ ID NO: 32:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 304 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-752-844-32

Query Match 39.6%; Score 282.8; DB 4; Length 304;
Best Local Similarity 97.6%; Pred. No. 4.7e-76;
Matches 287; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 358 caggtgcaggtgaaggagtcaggacccctggctggcgccctcacagagccctgtccatc 417
Db 1 CAGGTGCAGCTGGTGGAGTCAGACACCTGGCTGGTGGCGCCCTCACAGAGCCCTGTCCATC 60
Oy 418 actgtcactgtctctgggttttcattaaacaaattatggtgtacactgggttcgccagcct 477
Db 61 ACTTGCACCTGCTCTGGGTTTTCATTAAACAGCATGTGTTGTTACACTGGGTTGCCAGCCT 120
Oy 478 ccaggaagggtctggagtggctgggagtaattatggctggtggaagcacaattataat 537
Db 121 CCAGGAAGGGTCTGGAGTGGCTGGGAGTAATTATGGCTGGTGGAGCACAAATATAAT 180
Oy 538 tcggctcttatgtccagactgagcatcagcaaggacaactccaagagcccaagtttctcta 597
Db 181 TCGGCTCTCATGTCCAGACTGAGCATCAGCAAAACAACTCCAAGAGCCCAAGTTTCTTA 240
Oy 598 aaatgaacagtcgtcaaaactgatgacacagccatgtactactgtgccagtcgg 651
Db 241 AAATGAACAGTCTGCAAACTGATGACACAGCCATGTACTACTGTGCCAGAGGG 294

RESULT 7
US-08-752-844-3
; Sequence 3, Application US/08752844
; Patent No. 5935821
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Malaya
; APPLICANT: Foon, Kenneth A.
; APPLICANT: Chatterjee, Sunil K.
; TITLE OF INVENTION: MONOCLONAL ANTIBODY 1A7 AND USE FOR THE
; TITLE OF INVENTION: TREATMENT OF MELANOMA AND SMALL CELL CARCINOMA
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:

Qy 292 gggaccaagctggaaat 308
|||||
Db 301 GGGACCAAGCTGGAGAT 317

RESULT 14

```

US-08-463-298-2
; Sequence 2, Application US/08453298
; Patent No. 5824304
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C/O FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,298
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 11-NOV-
 FILING DATE: 11-NOV-
 APPLICATION NUMBER: 13-NOV-
 FILING DATE: 13-NOV-
 ATTORNEY/AGENT INFORMATION:
 NAME: Haley Jr., Jam
 REGISTRATION NUMBER:
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 596-9
 TELEFAX: (212) 596-9
 INFORMATION FOR SEQ ID NO:
 SEQUENCE CHARACTERISTIC:
 LENGTH: 318 base pair
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-463-298-2

Query Match 38.7%; Score 276.2; DB 3; Length 318;
Best Local Similarity 94.6%; Pred. NO. 4.6e-74;
Matches 300; Conservative 0; Mismatches 8; Indels 9

QY	1	agtagtgatgaccagactcccaaatctctgcttgatcagcaggagacagggttac	60
DB	1	AGTATTGTGATGACCCAGACTCCCAAAATCCTGTTGTTTCAGCAGGACAGGGTTAC	60
QY	61	ataactcgaagccagtcagagctgagtaatgtagtgcttggtaccacaagaagcca	120
DB	61	ATAACTGCAAGGCCAGTCAGAGTGAGCTAATGATGTAGCTTGGTACCACAGAAAGCCA	120
QY	121	gggcagctctccgaactcgtgtatactctgcatccaatcgctacactggagtcctgat	180
DB	121	GGSCAGTCTCTAACTGCTGTATATTATGCAATCCAAATCGCTACACTGAGTCCCTGAT	180
QY	181	cgcttcactggcagtggaatatgggacggatttcaactttcaccaatcagcactgacaggt	240
DB	181	CGCTTCACCTGGCAGTGGATGGACCGGATTTACATTTCCACATCAGCACTGTGAGGCT	240

Qy	241	gaagacatgcgaatttatctctcgcagattatagct-----cgctcgaagg	291
Db	241	GAGACCTGGCAGTTATTCTCTGCAGCAGATTATAGCTCCGTACACGTTTCGAGGG	300
Qy	292	ggaccacagctggaaat	308
Db	301	GGGACCAAGCTGGAGAT	317

RESULT 15

```

US-08-436-339A-2
; Sequence 2, Application US/08436339A
; Patent No. 5843438
; GENERAL INFORMATION:
; APPLICANT: Papayannopoulou, Thalia (USA only)
; APPLICANT: Board of Regents, U.
; APPLICANT: Washington (except USA)
; TITLE OF INVENTION: PERIPHERALIZATION OF HEMATOPOIETIC STEM
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: c/o FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,339A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,128
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: PCT/US93/11060
; FILING DATE: 11-11-1993
; APPLICATION NUMBER: US 07/977,702
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B173C1P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-436-339A-2

```

Query Match 38.7%; Score 276.2; DB 3; Length 318;
Best Local Similarity 94.6%; Pred. No. 4.6e-74;
Matches 300; Conservative 0; Mismatches 8; Indels 9; Gaps 1;

Qy	1	agtatgtgatgaccagactccaataatctctgtgtatcagcaggagacagggttacc	60
Db	1	AGTATTGTGATGACCCAGACTCCCAATTCCTGTTTTCAGCAGGACAGGGTTACC	60
Qy	61	ataacctcaaggccagtcagagtgagtgatgatgtggctgtgtacaaacaagaacca	120
Db	61	ATAACTCGAAGGCCAGTCAGAGTGTGACTATATGATGTAGCTTGTGTACCACAGAAAGCCA	120
Qy	121	gggcagctccgaaactcgtgatatactctgcatccaatcgctacactggagtcctgat	180

Db 121 GGGCAGTCTCCTAACTGCTGATATATATGTCATCCAAATCGCTACACTGGAGTCCCTGAT 180
Qy 181 cgcttcaactggcagtgatagggacggaatttcactttcaccatcagcactgtgcaggct 240
|||||
Db 181 CGCTTCACTGGCAGTGGATATGGACGGATTTCACTTTACCATCAGCACTGTGCAGGCT 240
Qy 241 gaagacctggcagtttatttctgtcagcaggattatagct-----cgctcggagg 291
|||||
Db 241 GAAGACCTGGCAGTTTATTCTGTGACAGGATTATAGCTCTCCGTACACGTTTCGGAGG 300
Qy 292 gggaccaagctggaaat 308
|||||
Db 301 GGGACCAAGCTGGAGAT 317

Search completed: November 30, 1999, 14:10:21
Job time: 5147 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - ,nucleic search, using sw model

Run. on: November 30, 1999, 14:09:30 ; Search time 545.98 Seconds
(without alignments)
5591.774 Million cell updates/sec

Title: US-08-940-544-4

Perfect score: 714

Sequence: 1 agtattgtgatgccacagac.....cctcagtcacogtctctca 714

Scoring table: IDENTITY_NUC

Searched: 780561 seqs, 2137953050 residues

Database : GenEmbl.*

1: gb_bal.*

2: gb_baz.*

3: gb_om.*

4: gb_ov.*

5: gb_pat.*

6: gb_ph.*

7: gb_pl1.*

8: gb_pl2.*

9: gb_pr1.*

10: gb_pr2.*

11: gb_pr3.*

12: gb_ro.*

13: gb_sts.*

14: gb_sy.*

15: gb_un.*

16: gb_vl.*

17: em_fun.*

18: em_hcg.*

19: em_humi.*

20: em_hum2.*

21: em_in.*

22: em_om.*

23: em_or.*

24: em_ov.*

25: em_pat.*

26: em_ph.*

27: em_pl.*

28: em_ro.*

29: em_sts.*

30: em_sy.*

31: em_un.*

32: em_vl.*

33: gb_btg1.*

34: gb_btg2.*

35: gb_in1.*

36: gb_in2.*

37: gb_bal.*

38: gb_ba2.*

39: gb_hum3.*

40: gb_hum4.*

41: gb_pr4.*

42: gb_btg3.*

43: gb_btg4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description

RESULT	1	MUSVKLI	MUSVKLI	738 bp	DNA	ROD	14-DEC-1995
LOCUS	1	MUSVKLI	Mus musculus (clone H11)	Ig kappa light chain V region (Vk), Ig heavy chain V region (Vh), and single chain Fv gene.			
DEFINITION	1	MUSVKLI	Mus musculus (clone H11)	Ig kappa light chain V region (Vk), Ig heavy chain V region (Vh), and single chain Fv gene.			
ACCESSION	1	L41689	L41689				
NID	1	G1119223	G1119223				
VERSION	1	L41689.1	GI:1119223				
KEYWORDS	1	Diels-Alder reaction; Ig heavy chain variable region; Ig kappa light chain; Ig light chain variable region; antigen H11; catalytic antibody; immunoglobulin.					
SOURCE	1	Mus musculus (strain NZB x BALB/c, sub-species musculus) (clone: H11) male adult splenocyte DNA.					
ORGANISM	1	Mus musculus musculus					
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1	1 (bases 1 to 738)					
TITLE	1	Suckling, C.J., Stimson, W.H., Proctor, G.R., Bence, L.H., Brooks, L., Khalaf, A.L. and Tedford, C.M.					
JOURNAL	1	Catalytic Antibodies - Designed and Accidental					
REFERENCE	1	J. Chem. Soc. Perkin Trans. 1 (22), 2777-2780 (1993)					
AUTHORS	1	2 (bases 1 to 738)					
	1	Brooks, L., Suckling, C.J. and Stimson, W.H.					

ALIGNMENTS

RESULT	1	MUSVKLI	MUSVKLI	738 bp	DNA	ROD	14-DEC-1995
LOCUS	1	MUSVKLI	Mus musculus (clone H11)	Ig kappa light chain V region (Vk), Ig heavy chain V region (Vh), and single chain Fv gene.			
DEFINITION	1	MUSVKLI	Mus musculus (clone H11)	Ig kappa light chain V region (Vk), Ig heavy chain V region (Vh), and single chain Fv gene.			
ACCESSION	1	L41689	L41689				
NID	1	G1119223	G1119223				
VERSION	1	L41689.1	GI:1119223				
KEYWORDS	1	Diels-Alder reaction; Ig heavy chain variable region; Ig kappa light chain; Ig light chain variable region; antigen H11; catalytic antibody; immunoglobulin.					
SOURCE	1	Mus musculus (strain NZB x BALB/c, sub-species musculus) (clone: H11) male adult splenocyte DNA.					
ORGANISM	1	Mus musculus musculus					
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1	1 (bases 1 to 738)					
TITLE	1	Suckling, C.J., Stimson, W.H., Proctor, G.R., Bence, L.H., Brooks, L., Khalaf, A.L. and Tedford, C.M.					
JOURNAL	1	Catalytic Antibodies - Designed and Accidental					
REFERENCE	1	J. Chem. Soc. Perkin Trans. 1 (22), 2777-2780 (1993)					
AUTHORS	1	2 (bases 1 to 738)					
	1	Brooks, L., Suckling, C.J. and Stimson, W.H.					


```

FEATURES             Location/Qualifiers
     source            1..840
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /note="recombinant gene"
     CDS                1..840
                        /codon_start=1
                        /product="IgM/kappa antibody"
                        /protein_id="CAB08854.1"
                        /db_xref="PID:e1167934"
                        /db_xref="PID:g2108309"
                        /db_xref="GI:2108309"
                        /translation="MAQVGDGIOMTQSPPLSVASAGEKVTMSCKSSQSVLYSSNQKNY
LAWQKQPSKLLIYWAISRESGVDRFTGSGSDFTLTISVQAEADLAVYICHQ
YLSLTFAGTKLEIKREGSKSGSEKLESEVNLVSGPGLVQPSLSITCVSG
FSLTSYGVHWVQSGKGLWGLVIRWGSTDYNAAFMSRLSITKDNKSQVFFRMS
LQADDTAIYCAKNYGYNDYWGQGTSTVTSAPKTTPGAAGKLLISEDLNDIK
DEL"
     misc_feature       19..354
                        /note="Light chain variable domain (VL) coding sequence"
     misc_feature       355..402
                        /note="Linker peptide coding sequence"
     misc_feature       403..759
                        /note="Heavy chain variable domain (Vh) coding sequence"
     misc_feature       760..774
                        /note="Heavy chain first constant domain (CH1) fragment
coding sequence"
     misc_feature       787..819
                        /note="c-myc tag coding sequence"
     misc_feature       826..837
                        /note="KDEL coding sequence"
     BASE COUNT        230 a 210 c 209 g 191 t
     ORIGIN
1 atgtgtatgaccagactcccaattctgtgtatcagcagagacagagttaccata 63
2  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3  22 ATCCAGATGACAGAGTCTCCCACTCTCTGCTGTCTCTGCAGGAGAAAGGTCACATG 81
4  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
5  64 acctgcaaggccagctcagagtggt-----gagtaaatgattgtggttgg 105
6  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7  82 AGCTGTAGTCCAGTCAAGAGTGTATATACAGTTCAAAATCAGAGAAGACTACTTGGCCTGG 141
8  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9  106 taccacagaagccaggcagctccgaactgctgtatatactatcgtcatccaatcgctac 165
10 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
11 142 TACCAGAGAAACACAGGCGAGTCCTCTAAACTGCTGATCTACTGGGCTCCACTAGGGAA 201
12 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13 166 actgagtcctcgtatccttcactgagtggtggtatgagcgggattcactttcaccatc 225
14 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
15 202 TCCGGTGTCTCTGATCGCTTCACAGGAGTGATCTGGACAGATTTTACTCTTACCATC 261
16 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 226 agcactgtcaggctgaagacctgagctgtatttctgtcagcaggattatagctcgc-- 284
18 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19 262 AGCAGTGTACAAGCTGAAGACCTGGCAGTTTATTACTGTCAATCAATACCTCTCCTCGCTC 321
20 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 284 ----tcggaggggggacacagctggaataaaagggtgagcggttcaggcgaggtggc 339
22 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 322 ACGTTCCGTCGGGGACCAAGCTGGAATAAAACGGGAGGTAATCCTCAGGATCTGGC 381
24 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 340 tctg---gcgggtgcgagtcgaggtcaggctgaagagtcaggacacctgctgtggcg 396
26 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
27 382 TCCGAATCCAAATCAGAGTCTGAGGTGAACCTGTGTGAATCTGACCTGGCCTAGTGCAG 441
28 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
29 397 ccctcacagagcctgtccatcactgtcactgctctctgtggttttcattaaacaaattatggt 456
30 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
31 442 CCCTCACAGAGCCCTGTCCATAACCTGCACAGCTCTCTGGTTCTCATTAACCTAGCTATGGT 501
32 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 457 gtacactgggttgcgagcctccaggaaagggtctgagtggtgggagtaatatgggct 516
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 502 GTACACTGGGTTCGCCAGTCTCCAGGAAGGGTCTCGAGTGGCTGGAGTGATATCGAGA 561
 QY 517 ggtggaagcacaaattataattcggctcttattgtccagagctgagcatcagcaggacaac 576
 Db 562 GGTGGAAGCACACACTACAATGCAGCTTTCATGTCCAGACTGAGCATCACCAGGACAAC 621
 QY 577 tcaagaagccaagttttcttaaaaaatgaacagctctcaaaactgatgacacagccatgtac 636
 Db 622 TCCAAGAGCCAAGTTTCTTTAAATGAACAGTCTCAAGCTGATGACACTGCCATATAC 681
 QY 637 tactgtgcagctcggggggtaactacgctgtatgttgactactgggtcgaagaac 696
 Db 682 TACTGTGCCAAAATTAATACTAGTAAGTACGGGGCTATGACTACTGGGCTCAAGGAACC 741
 QY 697 tcagtccacgtctcctca 714
 Db 742 TCAGTCACAGTCTCTGCA 759

RESULT 3
 SYN7A41
 LOCUS
 DEFINITION
 SYN7A41 743 bp DNA SYN 27-APR-1993
 Synthetic single-chain antigen-binding protein gene (7A4-1/212 SCA)
 constructed from the antigen-binding (Fab) fragment of mouse
 catalytic monoclonal antibody NPN43C9 and linker DNA, partial cds.
 M68968
 ACCESSION
 M68968
 NID
 9207839
 VERSION
 7A41/212 SCA protein; fusion protein; monoclonal antibody;
 KEYWORDS
 single-chain antigen-binding protein.
 SOURCE
 Mouse DNA and synthetic DNA.
 ORGANISM
 synthetic construct
 artificial sequence.
 REFERENCE
 1 (bases 1 to 743)
 AUTHORS
 Gibbs, R.A., Posner, B.A., Filpula, D.R., Dodd, S.W., Finkelman, M.A.,
 J., Lee, T.K., Wroble, M., Whitlow, M. and Benkovic, S.J.
 TITLE
 Construction and characterization of a single-chain catalytic
 antibody
 JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 88, 4001-4004 (1991)
 MEDLINE
 91219500
 FEATURES
 Location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..735
 /gene="7A4-1/212"
 <1..735
 /gene="7A4-1/212"
 /codon_start=1
 /transl_table=11
 /product="7A4-1/212 SCA fusion protein"
 /protein_id="AAA73053.1"
 /db_xref="PID:g207840"
 /db_xref="GI:207840"
 /translation="DVMVTPSSLAWSGVQKVTMSCKSSQSLNISNKNYLAWVQO
 KQSPKLLIYFASTRSGVDRFTGSGSDFTLTISVQAEADQADYFCQOHYRAPR
 TFGGCTKLEIKGSTSGSKSGKQGLVSGPGLVAPQSLSITCTVSGTSLRYN
 VHWVQSPGKGLWGLMIMWGGSIENPALKSRLSITKDNKSQIFLKNLSLTQDDSA
 MYCYSGYGGGRFYSYWGQGLTVTS"
 misc_feature 1..339
 /gene="7A4-1/212"
 /note="light-chain variable region of mouse antibody
 NPN43C9"
 340..381
 misc_feature /gene="7A4-1/212"
 /note="synthetic linker"
 382..735
 misc_feature /gene="7A4-1/212"
 /note="heavy-chain variable region of mouse antibody
 NPN43C9"
 195 a 176 c 190 g 182 t

Query Match 58.4%; Score 417.2; DB 12; Length 840;
 Best Local Similarity 76.3%; Pred. No. 1.9e-104;
 Matches 563; Conservative 0; Mismatches 148; Indels 27; Gaps 3;

13

FT mat_peptide 1. 738 /product='inhibitor of binding of IL-6 to its receptor'
FT receptor'
FEATURES Location/Qualifiers
Source 1. 738
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 187 a 189 c 180 g 182 t
ORIGIN

Query Match 55.5%; Score 396.2; DB 5; Length 738;
Best Local Similarity 75.5%; Pred. No. 1.le-98;
Matches 555; Conservative 0; Mismatches 153; Indels 27; Gaps 4;

Qy 4 attgtgatgacccagactccaaattctgtctgtatcagcagagagagaggttaccata 63
Db 7 ATCTGTGACACAGTCTCCAAATTCCTGTGTATCAGCAGGAGCAGGTTTACCATA 66
Qy 64 acctgcaaggccagtcagagtgtagtgatgagtggtgtgtacccaagagccagg 123
Db 67 ACCTCAAGGCCAGTCAGAGTGTGAGTACTGTATGTTGTACCAAGCCAGGG 126
Qy 124 cagtcctcgaactctgatatactctgcatactcgaatcgaactggtggtccctgatcgc 183
Db 127 CAGTCTCTAACTACTGATATATCTATGATCATCCAAATCGCTACACTGGAGTCCCTGATCGC 186
Qy 184 ttactggcagtggtatggagcagatttcactttccatccatcagcagctgacagctgaa 243
Db 187 TTCATCGCAGTGGATATGGACGAGATTTCATTTTCCATCAGCAGTGTGCGAGCTGAA 246
Qy 244 gacctggcagttattctgtcagcagagattatag-----ctcgtctggagggggg 294
Db 247 GACCTGGCAGTTTATTTCTGTGACGAGGATTATAGTCTCCATTCACGTTGCGTCGGG 306
Qy 295 accaagctggaataaaagggtgagggcgttcagcgcgaggtggtctcggcgtgagcga 354
Db 307 ACAAAGTTGGAATAAATA---GTCGACAAATCCTCAGGATCTGGCTCGGAATCCAAAGC 363
Qy 355 tcgacagtcaggtgaaagggtctgagtggtcagcgcgtggtggtcgtggtgagcga 468
Db 364 ACAGAGTCAAACTCGAGAGTCTGGCCCTGGGATATTCAGCCCTCCAGACCTCAGT 423
Qy 415 atcacttgactgtctctgggttttcattaaaccaattat-----ggtgtacactgggtt 468
Db 424 CTGACTGTGTTCTTCTCTGGGTTTCTACTGAGCACTTCTGTTATGGGTGTGAGTGGATT 483
Qy 469 ggcagcctcaggaaggggtctgtgagtggtggtggtggtggtggtggtggtggtggtggt 528
Db 484 COTCAGCCTTCAGGGAAGGGTCTGGAGTGGGTGGGACACATTTATTTGGGATGATGACAAA 543
Qy 529 aattataattcgtctcttattgtccagactgagcagcagcagcagcagcagcagcagcagc 588
Db 544 CACTATACCCATCCCTGAAGCGGCTCAAAATCTCCAGGATACCTCCACCAACAGC 603
Qy 589 gttttcttaaaataaagcagtgatgacacagcagcagcagcagcagcagcagcagcagcagc 648
Db 604 GTATTCTCAAGATCACCAGTGTGACACTGCAGATACCTGCCACATCTTCTGTCTCGA 663
Qy 649 cg-----gggggggtaactacggctatgctttggactactgggggtcaaggaacctca 699
Db 664 AGAATCTCTATGTTGGGGGACATGCTATGGACTATGCTATGGGCTCAAGGAACCTCA 723
Qy 700 gtcacccgtctctca 714
Db 724 GTACCCGTCTCTCA 738

RESULT 7
LOCUS I45910 738 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 6 from patent US 5639455.
ACCESSION I45910

NID 92469875
VERSION 145910.1 GI:2469875
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 738)
AUTHORS Shimamura,T., Nakazawa,H. and Hamuro,J.
TITLE Immunosuppressant
JOURNAL Patent: US 5639455-A 6 17-JUN-1997;
FEATURES Location/Qualifiers
source 1. 738
/organism='unknown'
BASE COUNT 187 a 189 c 180 g 182 t
ORIGIN

Query Match 55.5%; Score 396.2; DB 5; Length 738;
Best Local Similarity 75.5%; Pred. No. 1.le-98;
Matches 555; Conservative 0; Mismatches 153; Indels 27; Gaps 4;

Qy 4 attgtgatgacccagactccaaattctgtctgtatcagcagagagagaggttaccata 63
Db 7 ATCTGTGACACAGTCTCCAAATTCCTGTGTATCAGCAGGAGCAGGTTTACCATA 66
Qy 64 acctgcaaggccagtcagagtgtagtgatgagtggtgtgtacccaagagccagg 123
Db 67 ACCTCAAGGCCAGTCAGAGTGTGAGTACTGTATGTTGTACCAAGCCAGGG 126
Qy 124 cagtcctcgaactctgatatactctgcatactcgaatcgaactggtggtccctgatcgc 183
Db 127 CAGTCTCTAACTACTGATATATCTATGATCATCCAAATCGCTACACTGGAGTCCCTGATCGC 186
Qy 184 ttactggcagtggtatggagcagatttcactttccatccatcagcagctgacagctgaa 243
Db 187 TTCATCGCAGTGGATATGGACGAGATTTCATTTTCCATCAGCAGTGTGCGAGCTGAA 246
Qy 244 gacctggcagttattctgtcagcagagattatag-----ctcgtctggagggggg 294
Db 247 GACCTGGCAGTTTATTTCTGTGACGAGGATTATAGTCTCCATTCACGTTGCGTCGGG 306
Qy 295 accaagctggaataaaagggtgagggcgttcagcgcgaggtggtcgtggtgagcga 354
Db 307 ACAAAGTTGGAATAAATA---GTCGACAAATCCTCAGGATCTGGCTCGGAATCCAAAGC 363
Qy 355 tcgacagtcaggtgaaagggtctgagtggtcagcgcgtggtggtcgtggtgagcga 468
Db 364 ACAGAGTCAAACTCGAGAGTCTGGCCCTGGGATATTCAGCCCTCCAGACCTCAGT 423
Qy 415 atcacttgactgtctctgggttttcattaaaccaattat-----ggtgtacactgggtt 468
Db 424 CTGACTGTGTTCTTCTCTGGGTTTCTACTGAGCACTTCTGTTATGGGTGTGAGTGGATT 483
Qy 469 ggcagcctcaggaaggggtctgtgagtggtggtggtggtggtggtggtggtggtggtggt 528
Db 484 COTCAGCCTTCAGGGAAGGGTCTGGAGTGGGTGGGACACATTTATTTGGGATGATGACAAA 543
Qy 529 aattataattcgtctcttattgtccagactgagcagcagcagcagcagcagcagcagcagc 588
Db 544 CACTATACCCATCCCTGAAGCGGCTCAAAATCTCCAGGATACCTCCACCAACAGC 603
Qy 589 gttttcttaaaataaagcagtgatgacacagcagcagcagcagcagcagcagcagcagcagc 648
Db 604 GTATTCTCAAGATCACCAGTGTGACACTGCAGATACCTGCCACATCTTCTGTCTCGA 663
Qy 649 cg-----gggggggtaactacggctatgctttggactactgggggtcaaggaacctca 699
Db 664 AGAATCTCTATGTTGGGGGACATGCTATGGACTATGCTATGGGCTCAAGGAACCTCA 723
Qy 700 gtcacccgtctctca 714
Db 724 GTACCCGTCTCTCA 738

```

RESULT 8
MUSALCA
LOCUS      855 bp      mRNA      ROD      12-JUL-1995
DEFINITION Mus musculus (clone pCT.kvhd1) antibody mRNA, 3' end of cds.
ACCESSION  L43544
NID        g896291
VERSION    L43544.1  GI:896291
KEYWORDS   antibody.
SOURCE     Mus musculus cDNA to mRNA.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 855)
AUTHORS    Paul,S., Sun,M. and Gao,Q.-S.
TITLE      Catalysis by natural and hybrid single chain Fv constructs
JOURNAL    Unpublished (1995)
FEATURES   Location/Qualifiers
            source          1..855
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /cell_type="hybridoma"
                        <!--.354
                        /note="putative"
                        /product="antibody light chain"
                        <!--.855
                        /codon_start=1
                        /product="antibody"
                        /protein_id="AAA69766.1"
                        /db_xref="PID:g896294"
                        /db_xref="GI:896294"
                        /translation="DVLMTQPLTSLVITGQPASISCKSSQLLHTDGKTYLIWLLQR
                        PGSPKRLIYLSKLDSPDRFTGSGSDTFLKISVREADLGVYVCWQTHFPQT
                        FGGTKLEIKRADAPGTSKSGKSEKQVQLQESGPGLVAPQSLSITCTVSGFS
                        LTGYGVNVRQPGKGLWLMIGWDGNTDYNALKSRISLISKDNKSQVFLKMNLSLH
                        TDDTARYICARERDYRLDYWGQGTITVVSSTTKTPPSVYPAHAHHHHGAEQKLIS
                        EEDLNGAA"
                        <!--.354
                        /note="This CDS feature is included to show the
                        translation of the corresponding V-region. Presently
                        translation qualifiers on V_region features are illegal."
                        /codon_start=1
                        /protein_id="AAA69764.1"
                        /db_xref="PID:g896292"
                        /db_xref="GI:896292"
                        /translation="DVLMTQPLTSLVITGQPASISCKSSQLLHTDGKTYLIWLLQR
                        PGSPKRLIYLSKLDSPDRFTGSGSDTFLKISVREADLGVYVCWQTHFPQT
                        FGGTKLEIKRADAP"
                        355..396
                        /note="putative"
                        /function="linkers between VL and VH domains"
                        <397..744
                        /note="putative"
                        /product="antibody heavy chain"
                        <397..>744
                        /note="This CDS feature is included to show the
                        translation of the corresponding V-region. Presently
                        translation qualifiers on V_region features are illegal."
                        /codon_start=1
                        /protein_id="AAA69765.1"
                        /db_xref="PID:g896293"
                        /db_xref="GI:896293"
                        /translation="QVQLQESGPGLVAPQSLSITCTVSGFSLTGYGVNVRQPPGKG
                        LEWLGMIGWDGNTDYNALKSRISLISKDNKSQVFLKMNLSLHTDDTARYICARERDYR
                        LDYWGQGTITVVS"
                        745..774
                        /note="c-myc peptide-1; putative"
                        /function="facilitates recombinant protein identification"
                        784..801
                        /note="poly-histidine; putative"
                        /function="facilitates recombinant protein purification"
                        231 a 217 c 217 g 190 t

BASE COUNT 231 a 217 c 217 g 190 t
ORIGIN

```

```

Query Match      49.4%; Score 352.6; DB 12; Length 855;
Best Local Similarity 71.8%; Pred. No. 1e-86;
Matches 539; Conservative 0; Mismatches 164; Indels 48; Gaps 4;

QY 3 tattgtgagcaccagactcccaaatccctgctgttatcagcaggagacagaggttaccat 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 TGTGTTGATGACCCAACTCCACTCACTTGTGCGGTATACCATTTGGACAAACCAAGCATCAT 62

QY 63 aacctcaaggccagtcagag-----tgtagtaaatgatgtgctgtgga 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 CTTCTGCAGTCAAGTCAGAGGCTCTTACATACTGATGGAAGACATATTTGATTGGTT 122

QY 108 ccaacagaagcaggcagctctccgaaactgctgatactctgcatcctcaactcgtctac 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 GTTACAGAGCCAGGCCAGCTCTCCAAAGCCCTAATCTATCTGGTGTCTAAACTGGACTC 182

QY 168 tggagtcctctgctccttcaactggcagtgagtatatgggaaggatttcacattccaccat 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 TGGAGTCCCTGACAGGTTCACTGGCAGTGGATCAGGACAGATTTTCACACTGAAATCAG 242

QY 228 cactgtgcaggctgaagacctggtcagtttattctgtcagcagg-----attatag 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 CAGAGTGGAGGCTGAGGATTTGGAGTTTATTATGCTGGCAAGGTACACATTTTCCCTCA 302

QY 279 ctgcctcggaggggggaccagctggaaataaaagtgagcgggttcaggcgagggtg 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 GACGTTGGTGGAGGCCACCAAGCTGGAATCAACGGGGTGATGCTGCACACAGGTTCTAC 362

QY 339 ctctggcgggtgcgggagtcg-----caggtgcagggtgaaggagtcaggacc 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CTCGGGCTCTGGGAAATCGTCTCAAGGTAAAGGACAGGTGCAGCTGCAGGATCAGGAC 422

QY 384 tggcctgggtggcgcctcacagacgctgccatcactgctcctctctgtggttttcatt 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 TGGCCTGGTGGCGCCTCACAGAGGCTGCCATCATGACCCTCAGGCTCTCAGGTTTCTCAT 482

QY 444 aaccaattatggtgtacctggttcgcagcctccagaaagggtctggagtgcctg 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 AACCGGCTATGTTGTAACATGGGTTTCGCGCAGCTCCAGGAAAGGCTCTGGAGTGGCTGG 542

QY 504 agtaatatgggctggtggaagcacaaattataatcggctcttatgtccagactagcat 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 AATGATTTGGGTGATGGAACACACAGACTATAATTCAGCTCTCAAAATCCAGACTCAGCAT 602

QY 564 cagcaggcacaactccaagagcagaagtttcttaaaatgaacagctcgcacactgatga 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 CAGCAGGACAACTCCAGAGCCAAAGTTTCTTAAAAATGAACAGTCTGCACACTGATGA 662

QY 624 cacagcatgtactactgtgccagtcggggggttaactacggtctatgctttggactactg 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 CACAGCCAGGTACTACTGTGCCAGAGAGAGAT-----TATAGGCTGACTACTG 713

QY 684 gggtaagaagacctcagtcacccgtctctc 714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 714 GGGCCAAGSCACCGGTCAACGCTCTCCTCA 744

RESULT 9
SYNCHIPL
LOCUS      1701 bp      DNA      SYN      29-OCT-1994
DEFINITION Cloning vector encoding chimeric plasminogen activator K12G0S32
            gene, complete cds.
ACCESSION  M69273
NID        g208116
VERSION    M69273.1  GI:208116
KEYWORDS   chimeric plasminogen activator; secreted protein; synthetic gene.
SOURCE     Cloning vector DNA.
ORGANISM   unidentified cloning vector
            artificial sequence; vectors.
REFERENCE  1 (bases 1 to 1701)
AUTHORS    Holvoet,P., Laroche,Y., Lijnen,H.R., Van Cauwenberge,R.,

```


/transl_table=11
/protein_id="AA73072.1"
/db_xref="GI:208808"
/translation="MRTPAQLGILLWFFGKIDKIMQSPSSMYASLGERVTVCCK
ASODINSYLSWIOQPKPKPTLIYRGNRLVAGVRFSGSGGQDYSLTISLSEYED
GVVYCLRYDERPFTFGSTKLEIKRAGOGSSVOVOLKSGPLVOPQSLSITCTVVS
GSLITYGVHVIHQSPGKGLWGLVWISGSDYNAFISRLSINKDNSKQSVFFKMN
SLQANDIAYIYCARNTWGTSMDFWGGTISVTSS"
mat_peptide 215 a 188 c 202 g 219 t
BASE COUNT 72 .767
ORIGIN

Query Match 49.2%; Score 351; DB 14; Length 824;
Best Local Similarity 70.7%; Pred. No. 2.8e-86;
Matches 503; Conservative 0; Mismatches 190; Indels 18; Gaps 2;
QY 4 attgtgatgaccagactcccaattcctgttattatcagcaggagacaggtttaccata 63
DB 75 ATCAAGATGACCCAGTCTCCATCTTCATGATGATGATCTTAGGAGAGAGTCACTGTC 134
QY 64 acctgaagccagctcagagtgtagtaagtgtgctgtgtaccacagaaagccagg 123
DB 135 ACITGCAAGCGGAGTACAGACATTAATAGCTATTAGCTGGATCCAGCAGAACACAGG 194
QY 124 cagtctccgaactgctgatatactctgtcctcctacactgagtgagtcctctgacgc 183
DB 195 AAATCTCTTAAGACCTGATCTACCGTGGAAACAGATTGTTGCTGGGGTCCCATCAAG 254
QY 184 ttaactgagcagtgatgagcaggtattcaacttccaccatcagcactgtgcaggtgaa 243
DB 255 TTCAGTGGCAGTGGATGGGCAAGATTATCTCTCACCACAGCAGCCTGGAGTATGAA 314
QY 244 gacctgcagttattctctcagcagattatagctcgtcgtcggagggggacacactg 303
DB 315 GATGTGGAGTTATTATTGTTACGGTATGATGAGTTTCATTCAGTTTCGGGTCGGGG 374
QY 304 gaataaaagtgagagcgttcaggcggaggtgctggtgagtggtgagtgagtgagtg 363
DB 375 ACAAGTGTGAAATCAAGCGTGGTCAAGGCTCTCTG-----TTCAAGTT 422
QY 364 caggtgaagagtcagcagactgctggtggtggtggtggtggtggtggtggtggtggtg 423
DB 423 CAGTGAAGCAGTCAAGCAGTGGCTGATGAGCGCTGATGAGCGCTGATGAGCGCTGATG 482
QY 424 actgtctctgggttttctatcaacaaattatggtggtggtggtggtggtggtggtggtg 483
DB 483 ACAGTCTCTGGTTCTCATTAACCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 542
QY 484 aagggctgagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 543
DB 543 AAGGCTGTGAGTGGCTGGGAGTATATGAGTGGTGGAGGAGCAGACACTATAATGCACT 602
QY 544 ctatgtccagactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 603
DB 603 TTCATATCCAGACTGAGCATCAACAAGGACAAATTCACAGAGCCAAAGTTTCTTTAAATG 662
QY 604 aacagctcgaactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 663
DB 663 AACAGTGTGAAGTAAATGACAGCCATATATATCTGTTGTTGTTGTTGTTGTTGTTGTTG 716
QY 664 ggcctgtcttgactactggtggtggtggtggtggtggtggtggtggtggtggtggtggt 714
DB 717 GGAACCTCTATGACTACTTGGGGTCAAGGAACCTCAGTCCACCGTCTCTCTCA 767

RESULT 11
MUSIGHAEI 354 bp mRNA ROD 27-APR-1993
LOCUS
DEFINITION Mouse Ig heavy-chain mRNA V region, partial cds from hybridoma H61-15VH.
ACCESSION M36228

g194839
M36228.1 GI:194839
V-region: immunoglobulin heavy chain; processed gene.
SOURCE Mouse (strain BALB/c), cDNA to mRNA, from hybridoma H61-15.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 354)
AUTHORS Kavalier J., Caton A.J., Staudt L.M., Schwartz D. and Gerhard W.
TITLE A set of closely related antibodies dominates the primary antibody response to the antigenic site CB of the A/PR/8/34 influenza virus hemagglutinin
J. Immunol. 145, 2312-2321 (1990)
JOURNAL 90375932
MEDLINE
COMMENT Draft entry and computer-readable sequence for [J. Immunol. (1990) In press] kindly submitted by J.Kavalier, 06-JUL-1990.
FEATURES
Location/Qualifiers
source 1..354
/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_line="H61-15"
/tissue_type="hybridoma"
/map="chromosome 12"
1..354
/gene="IgH"
CDS <1..>354
/gene="IgH"
/codon_start=1
/product="immunoglobulin heavy chain V-region"
/protein_id="AA38067.1"
/db_xref="PID:g194840"
/db_xref="GI:194840"
/translation="QVQLKESGPGLVAPQSLSITCTVSGFSLTSGVHWYRPPGKG
LEWGLVWAGGTNTNLSRLSKNSQVFLKMSLQTDITAMTCARGHYG
YYAMDYWGQGTSTVTS"
BASE COUNT 90 a 89 c 93 g 82 t
ORIGIN

Query Match 45.3%; Score 323.6; DB 12; Length 354;
Best Local Similarity 94.6%; Pred. No. 9.8e-79;
Matches 335; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 358 caggtcaggtgaagagtcagcagcctggtggtggtggtggtggtggtggtggtggtggt 417
DB 1 CAGGTGACCTGAAGGAGTCAGGACCTGGCTGGTGGCGCCCTCAGAGCCTGTCCATC 60
QY 418 acttgcaactgtctctgggttttctatcaacaaattatggtgacactgggttcgcccagct 477
DB 61 ACTTGACATGCTCTGGGTTTCATTAAACGAGTATGGTGTACACTGGGTTCGCCAGCCT 120
QY 478 ccaggaagaggtcgtgagtggtggtggtggtggtggtggtggtggtggtggtggtggt 537
DB 121 CCAGGAAGAGGTCTGGAGTGGCTGGGAGTAATATGGCTGGTGGAGAACACAAATTAAT 180
QY 538 tcggtcttattgcccagactgagcagcagcagcagcagcagcagcagcagcagcagcagc 597
DB 181 TCGGCTCTCATGTCCAGACTGAGCATCAGCAAGAACAACTCCCAAGAGCCAAAGTTTCTTA 240
QY 598 aaaaagaagctgtgaaactgagcagcagcagcagcagcagcagcagcagcagcagcagc 657
DB 241 AAATGAACAGTCTGCAACTGATGACAGCCATGCTACTGTGCCAGAGGGCATTC 300
QY 658 aactacggtctgttggactactggtggtggtggtggtggtggtggtggtggtggtggt 711
DB 301 TACGGTACTATGCTATGGACTACTGGGTCAAGGAACCTCAGTCCACCGTCTCC 354
RESULT 12
MUSIGHQ520

LOCUS MUSIGM0520 345 bp mRNA ROD 29-OCT-1994
 DEFINITION Mouse IgM chain mRNA VJ4 region, family Q52N.
 ACCESSION M76414
 NID 9197943
 VERSION M76414.1 GI:197943
 KEYWORDS J-region: V-region: immunoglobulin heavy chain; immunoglobulin-kappa.
 SOURCE Mus musculus (strain BALB/c, sub_species domesticus) female ascites CDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 345)
 AUTHORS Deng, Y.J., Chua, M.M., Andrews, G.C. and Karush, F.
 TITLE Primary B-cell response to neuropeptide Y and bovine pancreatic polypeptide
 JOURNAL Mol. Immunol. 29 (7-8), 847-856 (1992)
 MEDLINE 92342150
 FEATURES
 source
 Location/Qualifiers
 1..345
 /organism="Mus musculus"
 /strain="BALB/c"
 /sub_species="domesticus"
 /db_xref="taxon:10090"
 /cell_line="DSP2.2"
 /cell_type="hybridoma"
 /sex="female"
 /tissue_type="ascites"
 /map="chromosome 12"
 1..345
 /gene="IgM"
 1..345
 /partial
 /gene="IgM"
 /codon_start=1
 /product="mu-immunoglobulin"
 /protein_id="AA51159.1"
 /db_xref="PID:g197944"
 /db_xref="GI:197944"
 /translation="LESGPGLVAPQSLSITCTVSGFSLTSYGVHWVRQPPGKGLWLVGIWAGSTNTNSALMSRLSISKNSKQVFLKNSLQTDITAMYCASPLITYYAMDYWGQTSVTYS"
 4..278

gene

CDS

V_region

N_region

D_segment

N_region

J_segment

BASE COUNT

ORIGIN

Query Match 44.1%; Score 314.8; DB 12; Length 345;
 Best Local Similarity 95.0%; Pred. No. 2.6e-76;
 Matches 325; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 373 gagtcaggacctggctggtggcgccctcacagagcgttccatcactgtcactgtctct 432
 |||||

Db 4 GAGTCAGGACCTGGCCCTGGTGGCGCCCTCACAGAGCCTGCCATCACTTGCACCTGCTCT 63
 |||||

QY 433 ggggtttcattacaattatggtgfaactgggttcgcagccctccagaaggggtctg 492
 |||||

Db 64 GGGTTTTCATTACCAAGCATATGTTGTACACTGGGTTTCGCCAGCCTCCAGGAAGGCTCTG 123
 |||||

QY 493 gagtggctggagtaataatgggctgggtggaagcacaaataataatcggtctttatgtcc 552
 |||||

Db 124 GAGTGGCTGGGAGTATATGGCTGGTGTGAGCAGACAAATATTAATTCGCTCTCATGTC 183
 |||||

QY 553 agactgagcatcagcaaggacaaactccaagcgcccaagtttttcttaaaatgaacagtctg 612
 |||||

Db 184 AGACTGAGCATCAGCAAAAGACAACCTCCAGAGCCCAAGTTTCTTAAAAATGAACAGTCTG 243
 |||||

QY 613 caaactgatgacacagccatgtactactgtccactcggggggtaactacagctatct 672
 |||||

Db 244 CAAACTGATGACACAGCCATGTACTGTGCCACCCCTTGATTACTATGCT 303
 |||||

QY 673 ttggactactgggtgcaaggaaacctcagtcaccgtctctctca 714
 |||||

Db 304 ATGGACTACTGGGTCAGGAACCTCAGTCACCGTCTCTCTCA 345
 |||||

RESULT 13
 MUSIGHADX
 LOCUS
 DEFINITION Mouse Ig heavy-chain mRNA V region, partial cds. H220-22VH.
 ACCESSION M36217
 NID 9194815
 VERSION M36217.1 GI:194815
 KEYWORDS V-region; immunoglobulin heavy chain; processed gene.
 SOURCE Mouse (strain BALB/c), CDNA to mRNA, from hybridoma H220-22.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 360)
 AUTHORS Kavalier, J., Caton, A.J., Staudt, L.M., Schwartz, D. and Gerhard, W.
 TITLE A set of closely related antibodies dominates the primary antibody response to the antigenic site CB of the A/PR/8/34 influenza virus hemagglutinin

JOURNAL J. Immunol. 145, 2312-2321 (1990)
 MEDLINE 90375932

COMMENT Draft entry and computer-readable sequence for [J. Immunol. (1990) In press] kindly submitted

by J.Kavaler, 06-JUL-1990.
 FEATURES
 source
 Location/Qualifiers
 1..360
 /organism="Mus musculus"
 /strain="BALB/c"
 /sub_species="domesticus"
 /db_xref="taxon:10090"
 /cell_line="H220-22"
 /tissue_type="hybridoma"
 /map="chromosome 12"
 1..360
 /gene="IGH"
 /x1->360
 /gene="IGH"
 /codon_start=1
 /product="immunoglobulin heavy chain V-region"
 /protein_id="AAA38056.1"
 /db_xref="PID:g194816"
 /db_xref="GI:194816"
 /translation="QVQLKESGPGCLVAPQSLSITCTVSGFSLTSYGVHWVRQPPGKGLWLVGIWAGSTNTNSALMSRLSISKNSKQVFLKNSLQTDITAMYCARGHYGDGYAMDYWGQTSVTYS"

BASE COUNT 91 a 89 c 96 g 84 t
 ORIGIN

Query Match 43.5%; Score 310.8; DB 12; Length 360;
 Best Local Similarity 93.6%; Pred. No. 3.2e-75;
 Matches 337; Conservative 0; Mismatches 17; Indels 6; Gaps 1;

QY 358 caggtgcagggtgaaggactcaggacctggctggtggccctcacagagctgtccatc 417
 |||||

Db 1 CAGGTGCAGCTGAAGAGTCAGGACCTGGCTGGTGGCGCCCTCACAGAGCTGTGCCATC 60
 |||||

QY 418 acttgacctgtctctgggttttctcattacaacattatggttacactgggttcgcagcct 477
 |||||

Db 61 ACTTGCACCTGTCTCTGGGTTTTCATTAAACAGCATGTGTGTACACTGGTTCGCCAGCCT 120
 |||||

QY 478 ccaggaaaagggtctggagtggtggtggaataatagggtggtggaagcacaaataataat 537
 |||||

[illegible]

Db	293	GCCAAAGTTCCTTTAAAAATGAACAGCTGTGCAAAATTGATGACACAGCCATATACTACTGTG	35
Qy	644	cca---gcgggggggtaactacgcgctatgcttggactactcgggtcgaagaacctcag	700
Db	353	CCAAAGCAGCTGGAACGAATCTTTTACTATGCTATGACTATTTGGGGTCAAGGAACCTCAG	412
Qy	701	tcaccgtctctctca	714
Db	413	TCACCGTCTCCTCA	426
RESULT	15		
MUSIGHCVX			
LOCUS	MUSTGHCVX	420 bp	mrna
DEFINITION	Mouse immunoglobulin heavy chain variable region (Igh-V) mRNA, end.		02-MAY-1994
ACCESSION	L31403		
NID	9476717		
VERSION	L31403.1	GI:476717	
KEYWORDS	immunoglobulin heavy chain; variable region.		
SOURCE	Mus musculus (strain A/J, sub-species domesticus) hyperimmunized spleen cDNA to mRNA.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Jeffrey,P.D., Schildbach,J.F., Chang,C.Y., Kussie,P.H., Sheriff and Margolies,M.N.		
TITLE	Structure and specificity of the anti-digoxin antibody 40-50		
FEATURES	Unpublished (1994)		
	Location/Qualifiers		
	1..420		
	source		

```

TITLE      Structure and specificity of the anti-digoxin antibody 40-50
JOURNAL    Unpublished (1994)
FEATURES   Location/Qualifiers
            source          1..420
                        /organism="Mus musculus"
                        /strain="A/J"
                        /sub_species="domesticus"
                        /db_xref="taxon:10090"
                        /cell_line="4050"
                        /cell_type="hybridoma"
                        /tissue_type="hyperimmunized spleen"
                        /map="12"
sig_peptide 1..57
            /gene="Igh-v"
            1..>420
            /gene="Igh-v"
            /codon_start=1
            /product="immunoglobulin heavy chain variable region"
            /protein_id="AAA38191.1"
            /db_xref="PID:g476718"
            /db_xref="GI:476718"
            /translation="NAVLVFLCLVAFPSCVLSVHVKSGPLVAPQSLSITL
            GFSLITYGVHFFRPQPGKGLWGLIAGGNTDYNALMSRLSINKNKSQVIT
            SQAQDDATMYCYCARFFASYYDYADVYWGQGTSTVTTSS"
            1..420
            /gene="Igh-v"
            58..>420
            /gene="Igh-v"
            /product="Immunoglobulin heavy chain variable region"
BASE COUNT 94 a 112 c 102 g 112 t

```

```

gene
    /gene="igh-v"
mat_peptide
    /gene="igh-v"
    /product="Immunoglobulin heavy chain variable region"
BASE COUNT      94 a 112 c 102 g 112 t
SQADDTAMTYCARFRFASYDYAVDYWGQGTSTVSS"
1..420

```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 30, 1999, 12:44:32 ; Search time 383.27 Seconds
(without alignments)
4382.041 Million cell updates/sec

Title: US-08-940-544-3
Perfect score: 717
Sequence: 1 caggtgaactgcagcagtc.....cacagttggaataaaacgg 717

Scoring table: IDENTITY_NUC

Searched: 3032314 seqs, 1171202697 residues

Database: EST.*

- 1: em_est1:*
- 2: em_est2:*
- 3: em_est3:*
- 4: em_est4:*
- 5: em_est5:*
- 6: em_est6:*
- 7: em_est7:*
- 8: em_est8:*
- 9: em_est9:*
- 10: em_est10:*
- 11: em_est11:*
- 12: em_est12:*
- 13: em_est13:*
- 14: em_est14:*
- 15: em_est15:*
- 16: em_est16:*
- 17: em_est17:*
- 18: em_est18:*
- 19: em_est19:*
- 20: gb_est1:*
- 21: gb_est2:*
- 22: gb_est3:*
- 23: gb_est4:*
- 24: gb_est5:*
- 25: gb_est6:*
- 26: gb_est7:*
- 27: gb_est8:*
- 28: gb_est9:*
- 29: gb_est10:*
- 30: gb_est11:*
- 31: gb_est12:*
- 32: gb_est13:*
- 33: gb_est14:*
- 34: gb_est15:*
- 35: gb_est16:*
- 36: gb_est17:*
- 37: gb_est18:*
- 38: gb_est19:*
- 39: gb_est20:*
- 40: gb_est21:*
- 41: gb_est22:*
- 42: gb_est23:*
- 43: gb_est24:*
- 44: gb_est25:*
- 45: gb_est26:*
- 46: gb_est27:*
- 47: gb_est28:*
- 48: gb_est29:*
- 49: gb_est30:*
- 50: gb_est31:*
- 51: gb_est32:*
- 52: em_est20:*
- 53: em_est21:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	277	38.6	396	37	AA691311	AA691311 vs14f01.f
2	203.8	28.4	344	35	AA592800	AA592800 vo25g11.f
3	197	27.5	380	49	AI641928	AI641928 vq50c10.y
4	182.4	25.4	293	51	AV052316	AV052316 AV052316
5	182.2	25.4	310	43	AI180569	AI180569 uc70f09.f
6	180.2	25.1	461	43	AI201426	AI201426 q573c06.x
7	178.6	24.9	453	62	AI923980	AI923980 wh49d07.x
8	177.2	24.7	528	60	AI799757	AI799757 wc37g12.x
9	173.6	24.2	294	41	AI007196	AI007196 ua73g04.f
10	171.2	23.9	297	51	AV052647	AV052647 AV052647
11	170.6	23.8	488	60	AI791363	AI791363 ob68a09.y
12	159.8	22.3	387	35	AA569186	AA569186 nm30d10.s
13	126.6	17.7	393	20	T27593	T27593 EST100653.H
14	122.6	17.1	379	31	AA291381	AA291381 zt44g02.f
15	122.2	17.0	345	37	AA710291	AA710291 vt53a04.f
16	121.8	17.0	400	34	AA456778	AA456778 2v27f11.f
17	119.4	16.7	398	33	AA423447	AA423447 ve80a03.f
18	119.2	16.6	336	40	AA987559	AA987559 or83g01.s
19	119.2	16.6	470	44	AI306478	AI306478 qnl3b10.x
20	119.2	16.6	640	50	AI697483	AI697483 we14a04.x
21	119	16.6	413	31	AA301347	AA301347 EST14279
22	118.4	16.5	370	31	AA295093	AA295093 EST100400
23	117.2	16.3	431	37	AA710970	AA710970 vt93b09.f
24	115.6	16.1	328	31	AA295311	AA295311 EST10047
25	115.4	16.1	345	32	AA335086	AA335086 EST39457
26	114.6	16.0	402	48	AI549800	AI549800 ve80a03.y
27	114.4	16.0	279	31	AA300571	AA300571 EST113661
28	112.8	15.7	349	32	AA379044	AA379044 EST91999
29	112.6	15.7	471	43	AI233978	AI233978 EST230666
30	110	15.3	382	31	AA295786	AA295786 EST100987
31	110	15.3	333	31	AA300732	AA300732 EST13647
32	109.8	15.3	463	22	R69532	R69532 yj82d09.f1
33	109.6	15.3	420	42	AI120005	AI120005 uc25e04.f
34	109.4	15.3	253	31	AA295941	AA295941 EST101165
35	109	15.2	395	20	T29112	T29112 EST69384.Hu
36	109	15.2	335	31	AA318377	AA318377 EST20820
37	108.2	15.1	352	31	AA301261	AA301261 EST14181
38	107.8	15.0	363	32	AA367405	AA367405 EST78511
39	106.8	14.9	403	32	AA345486	AA345486 EST51505
40	106.6	14.9	291	32	AA377074	AA377074 EST89603
41	105.8	14.8	209	20	T28938	T28938 EST61186.Hu
42	104.4	14.6	287	20	T27609	T27609 EST101034.H
43	104.4	14.6	381	20	T29114	T29114 EST69430.Hu
44	104.4	14.6	823	29	AA170256	AA170256 ms87g10.f
45	104.2	14.5	383	20	T29916	T29916 EST99871.Hu

ALIGNMENTS


```

BASE COUNT      91 a      91 c      70 g      92 t
ORIGIN

Query Match      28.4%; Score 203.8; DB 35; Length 344;
Best Local Similarity 85.9%; Pred. No. 2.9e-48;
Matches 238; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 398 acatgagtcactcagtcagcaatcatgtctgcacccagggag--aaggtcac 455
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67 AAATGTTCTCACTCAATCCAGCAATCATGTCTGCTTCCAGAGGAGCAATGCTAC 126

QY 456 catgacctgcagtcgagctcaagtaagttatcatgcactggtaccagcagaagcctgt 515
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 CAGGACTGCAGTGCAGCTTCTACTGTAAGTTACATGCAATGTAACAGCAGAAGTCAG 186

QY 516 cactctcccaaaaagatgatttatgacacatccaaactggctctgtggagtcctctgcg 575
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 TACCTCCCCAAAACATGATTTATGATCATCTAACTGGCTTCTGGAATCCCTGCTCG 246

QY 576 cttagtgagtcggtctgggacctcttattctctcacaatcagcagcatggagcctgt 635
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 CTTCACTGGCAGTAGCTCTGGGACCTCTTACTCTCTCACAATCAGCAGAATTGAGGCTGA 306

QY 636 agatgtgcacttattactgcatcagcggagtagt 672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 ACATGTGCGCACTTATTACTGCGCAGCATTTGGAGTAGT 343

RESULT 3
AI641928      380 bp      mRNA      EST      29-APR-1999
LOCUS      vq50c10.y1 Barstead bowel MPLRB9 Mus musculus cDNA clone
DEFINITION IMAGE:1097682 5' similar to gb:M63550 Mouse Ig rearranged
            kappa-chain mRNA V-region, complete (MOUSE);, mRNA sequence.
ACCESSION  AI641928
NID         94720403
VERSION     1
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 380)
AUTHORS     Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
            Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
            Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
            Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
            Waterston, R. and Wilson, R.
            The WashU-NCI Mouse EST Project 1999
            Unpublished (1999)
TITLE       On Mar 10, 1998 this sequence version replaced gi:2948726.
JOURNAL
COMMENT

Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:603914
This read is a RESEQUENCE of a previously sequenced mouse clone
correct orientation)
This read has been verified (found to hit its original self in the
Seq primer: -40RP from Gibco
High quality sequence scop: 288.
Location/Qualifiers
1. .380
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1097682"

FEATURES
source
/clone_lib="Barstead bowel MPLRB9"
/tissue_type="bowel"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACGAATCTGAAGGGAGGAGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCCTGTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Source irradiated bowel harvested 72 hours after
irradiation (1400 Gys). Library constructed by Bob
Barstead."
BASE COUNT      107 a      92 c      84 g      96 t      1 others
ORIGIN

Query Match      27.5%; Score 197; DB 49; Length 380;
Best Local Similarity 84.7%; Pred. No. 2.7e-46;
Matches 232; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 390 cggatcgacatcagtcactcagtcctcagcaatcatgtctgcacccagggagaa 449
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 106 CAGAGGAGAAATGTCTCACCCAGTCTCCAGCAATCATGTCTGCACTCTAGGGAGAA 165

QY 450 ggtcaccatgacctgagtcgagtcgaagtataagttacatgcactggtaccagcagaa 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 GGTCAACATGAGTGCAGGGCCAGCTCAAGTGTAAATATACATGACTGTGTACCAGCAA 225

QY 510 gctctcactcccccacaaagatttatgacacatccaaactggctctgtgagtcacc 569
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 GTCAATGCTCTCCCAACATATGATTTATTACATATCCCAACCTGGCTCTGTGAGTCCC 285

QY 570 tctcgtctcagtcgagtcggtctggacctcttattctctcacaatcagcagcatgga 629
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 AGCTGCTTTCAGGGCAGTGGTCTGGGAACCTCTATTCTCTCACAAATCAGCATGNA 345

QY 630 ggctgtagatgtgcacttattactgcatcag 663
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 GGGTAAGATGCTGCGCAC-TATTACTGCGCAGCAG 378

RESULT 4
AV052316      293 bp      mRNA      EST      22-JUN-1999
LOCUS      AV052316 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
DEFINITION Clone 1810021J06, mRNA sequence.
ACCESSION  AV052316
NID         95138088
VERSION     AV052316.1 GI:5138088
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 293)
AUTHORS     Carinici, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
            Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
            Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
            Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H.,
            Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y.,
            Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y.,
            Tominaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T.,
            Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
            RIKEN Mouse ESTs
            Unpublished (1999)
TITLE       On Dec 4, 1997 this sequence version replaced gi:2570916.
JOURNAL
COMMENT

Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

```



```
BASE COUNT      85 a      68 c      77 g      64 t
ORIGIN

Query Match      24.2%; Score 173.6; DB 41; Length 294;
Best Local Similarity 78.6%; Pred. No. 1.2e-39;
Matches 231; Conservative 0; Mismatches 61; Indels 2; Gaps 2;

QY 3 ggtgaactcagcagtcaggacactgagc-tggtgnagcctgggcttcagtgaaagatat 61
Db 1 GGTCACAACTCAGCAGCCTGGGCTGAGCAGTGGTGAGGACTGGGTCTCAGTGAAGCTGT 60

QY 62 cctgcagaactcttggnacaaattcactgaatacacaccatgcactgggtgaagcagagcc 121
Db 61 ACTGCAAGGATTCAGGCTACACCTTCACCAGCTACTGGATGCATGGGTGAAGCAGAGGC 120

QY 122 atggaagagccttgagtgatggaggtatttaactcctaacaatggtggtactaaactaca 181
Db 121 ATATACAGGCATTGATGATGGTAAACATTGACCTTCAGCTATAGTGGAACTCACTACA 180

QY 182 agcagaattcgaagggcgaagccacattgactgactgacagagtcctccagcacagcctaca 241
Db 181 ATCAAAAGTTCAAGGACCAAGCCACATTGACTGTAGACAAATTCACCAGCACAGCCTACA 240

QY 242 -tgagctccgagcctgacatctgaggtctgcagctctattactgtgcaaga 294
Db 241 GTGAGCTCAGCAGACTGACATCTGAGGACTCTGCGGTCTATTACTGTGCAAGA 294

RESULT 10
AV052647      297 bp      mRNA      EST      22-JUN-1999
LOCUS      AV052647 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
DEFINITION      clone 1810022L19, mRNA sequence.
ACCESSION      AV052647
NID      95138419
VERSION      AV052647.1 GI:5138419
KEYWORDS
SOURCE      house mouse.
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 297)
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H.,
Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y.,
Sugahara, Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y.,
Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T.,
Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished (1999)
On Jun 22, 1998 this sequence version replaced gi:3246882.

Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtrc.riken.go.jp
Thermolabile and thermostable activation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtrc.riken.go.jp) for
further details.

FEATURES
Location/Qualifiers
1..297
/organism="Mus_musculus"
/strain="C57BL/6J"
```

```
/db_xref="taxon:10090"
/clone="1810022L19"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"

BASE COUNT      79 a      54 c      75 g      89 t
ORIGIN

Query Match      23.9%; Score 171.2; DB 51; Length 297;
Best Local Similarity 75.2%; Pred. No. 5.9e-39;
Matches 212; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 9 actgcagcagtcagcagcctgagcctgagcctgagcctgagcctgagcctgagcctgagcct 68
Db 16 AATGCAGCAGTTGGAGCTGAGGTGGTGACACCCCGGGCTTTAGTGAAGCTGCTTGCAAA 75

QY 69 gactctctganaacaattcactgaatacacaccatgcactgggtgaagcagagccatggaaa 128
Db 76 GGTCTTGGCTACACTTTCACTGAATATATATACACTGGGTAAAGCAGAGTTTGGACA 135

QY 129 gagccttgagtgagtgaggtatttaactcctaacaatggtggtactaaactacaagcagaa 188
Db 136 GGGCCTTGAGTGGATGGCTGCTGTTTTACCTGGAAAGTGGTAGTATAAATTACAAATGAGAA 195

QY 189 gttcaaggcgaagggcagcattgactgactgtagacaaagtcctccagcacagcctcatagagct 248
Db 196 TTTCAACGACGAAGGCCACATTGACTGCGGACAAATCTTCCAGCACAGTTTATATGAACT 255

QY 249 ccgcagcctgacatctgaggtgaggtctgcagctctattactgtgc 290
Db 256 TAGTAGATTGACATCAGAGAAATTTGCGGTCTATTTTGTGC 297

RESULT 11
AV052647      488 bp      mRNA      EST      02-JUL-1999
LOCUS      AV052647 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA
DEFINITION      clone 1810022L19, mRNA sequence.
ACCESSION      AV052647
NID      95339079
VERSION      AV052647.1 GI:5339079
KEYWORDS      human.
SOURCE      Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE      Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 488)
NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
On May 18, 1998 this sequence version replaced gi:3138279.
Other ESTs: oh68a09.x5
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
information
This 5' resequenced clone has no previous 5' data to verify this
new read against
```


GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 12:44:34 ; Search time 39.3 Seconds
(without alignments)
1929.363 Million cell updates/sec

Title: US-08-940-544-3
Perfect score: 717
Sequence: 1 cagggtgaactgcagcagtc.....cacagttggaataaaacgg 717

Scoring table: IDENTITY_NUC

Searched: 195662 seqs, 52875861 residues

D: se : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/3A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/3B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS9_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	521.6	72.7	726	3	US-08-553-497A-25
2	517.4	72.2	1797	2	US-08-463-163-2
3	512.6	71.5	726	3	US-08-553-497A-27
4	507.2	70.7	810	4	US-08-652-507-1
5	504.6	70.4	732	3	US-08-553-497A-19
6	503	70.2	732	3	US-08-553-497A-21
7	501.6	70.0	1679	3	US-08-661-052-15
8	487	67.9	738	3	US-08-553-497A-23
9	409.2	57.1	738	4	US-08-956-047-24
10	397	55.4	711	3	US-08-190-199A-64
11	386.8	53.9	708	3	US-08-190-199A-60
12	381	53.1	720	4	US-08-800-198-7
13	368.2	51.4	831	3	US-08-403-853-17
14	360.8	50.3	756	4	US-08-797-689-17
15	348.8	48.6	2012	1	US-08-235-838-15
16	348.8	48.6	2012	4	US-08-465-473B-15
17	347	48.4	748	1	US-08-235-838-10
18	347	48.4	748	4	US-08-465-473B-10
19	325.6	45.4	797	3	US-08-894-922A-13
20	325	45.2	672	3	US-08-190-199A-62
21	324	45.2	891	3	US-08-894-922A-9
22	314.8	43.9	669	3	US-08-190-199A-66
23	314.4	43.8	858	4	US-08-428-257A-71
24	299	41.7	2233	1	US-08-235-838-6
25	299	41.7	2233	4	US-08-235-838-13
26	299	41.7	2233	4	US-08-465-473B-6
27	299	41.7	2012	1	US-08-465-473B-13
28	298.8	41.7	748	4	US-08-235-838-4
29	298.8	41.7	748	4	US-08-465-473B-4
30	295.8	41.3	793	4	US-08-956-047-29
31	291.8	40.7	723	1	US-07-988-430-93
32	291.8	40.7	723	2	US-08-425-336-90
33	291.8	40.7	723	2	US-08-488-113B-90
34	291.8	40.7	723	2	US-08-477-484B-90
35	291.8	40.7	723	3	US-08-646-360-90
36	291.8	40.7	723	5	PT-US92-09487-93
37	280.6	39.1	354	3	US-08-491-845-5

Sequence 1, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 18, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 6, Appli

38 278.6 38.9 384 3 US-08-656-586-1
39 278.4 38.8 943 3 US-08-303-569B-4
40 278.4 38.8 943 4 US-08-116-247-4
41 276.8 38.6 1460 3 US-08-392-338A-18
42 276.2 35.5 2165 4 US-08-263-911-6
43 275.8 38.5 2165 4 US-08-263-911-8
44 275 38.4 309 2 US-08-467-393-3
45 272.2 38.0 780 3 US-08-447-402-6

ALIGNMENTS

RESULT 1
US-08-553-497A-25
; Sequence 25, Application US/08553497A
; Patent No. 5844093
; GENERAL INFORMATION:
; APPLICANT: KETLEBOROUGH, C. A.
; APPLICANT: BENDIG, MARY M.
; APPLICANT: ANSELL, KEITH H.
; APPLICANT: GUSSOW, DETLEF
; APPLICANT: ADAN, JAUME
; APPLICANT: MITJANS, FRANCES
; APPLICANT: ROSELL, ELISABET
; APPLICANT: BLASCO, FRANCESC
; APPLICANT: PIJLATS, JAUME
; TITLE OF INVENTION: ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553.497A
; FILING DATE: 17-NOV-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP95/00978
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94104160.0
; FILING DATE: 17-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94118970.6
; FILING DATE: 03-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HAMLET-KING, DIANA
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: MERCK 1726
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEC ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:

RESULT 2
US-08-463-163-2
; Sequence 2, Application US/08463163
; Patent No. 5696237

	Query Match	72.2%	Score 517.4;	DB 2;	Length 1797;
	Best Local Similarity	83.7%	Pred. No. 2.8e-137;		
	Matches 597;	Conservative	0;	Mismatches 113;	Indels 3; Gaps 1;
Qy	1	cagggtgaactgcagcagctcaggaccctgaactcgtgtgagcctcggggttcagtgaaagata	60		
Lb	4	CAGGTCCAGCTGCACGAGCTCTGGGGCTGAACGTGCCAAACCTGGGGCCCTCAGTGAAGATG	63		
Qy	61	tccctgcagaagactctctganaacaaattcactgaaTcacaccatgcactcgtgggtgaagcagagac	120		
Lb	64	TCCTGCAAGGCTTCTGGCTACACCTTTACTAGCTACAGGATGCATGGGTAAACACAGAGG	123		
Qy	121	catggaagaagcccttgagtggtatcggaggtattaatcctcaaatggtcgttactaactac	180		
Lb	124	CCTGGACAGGGCTCGGAATGGATGGATATATTAATCTAGCACTGGGTATCTACTGAATAC	183		

Db 482 ACAGCTCAAGTGTAGTTACATGTAAGTACAGGACAGGATCTCCCCGAC 541
Qy 530 gatggattatgacacatcccaactggcttctgagtgccctgctcgtcagtgccagtg 589
Db 542 TTCTGATTATGACACATCAACCTGGCTTCTGGAGTCCCTGTTGCTTCAAGTGGCAGTG 601
Qy 590 ggtctggagactcttattctctcaaatcagcagcatggaggtctgagatctccactt 649
Db 602 GGTCTGGAGACTCTTACTCTCTCAATCAGCCGAATGGAGGCTGAAGATCTGCCACTT 661
Qy 650 attactgcatcagcgggagtgattaccgctcagcttcgctgctgggacacagttgaaa 709
Db 662 ATTACTGCACAGTGGAGTAGTTACCCGCTACGTTCCGTTGCTGGGACCAAGCTGGAAA 721
Qy 710 taaa 714
Db 722 TAAAA 726

REMARK

US-08-652-507-1
; Sequence 1, Application US/08652507
; Patent No. 5876691
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: ANTIBODY AGAINST CARCINOEMBRYONIC ANTIGEN
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye, P.C.
; STREET: 1100 No. 5876691th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,507
; FILING DATE: 02-Jul-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-211
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-652-507-1

Query Match 70.7%; Score 507.2; DB 4; Length 810;
Best Local Similarity 82.2%; Pred. No. 1.6e-134;
Matches 597; Conservative 0; Mismatches 120; Indels 9; Gaps 1;

Qy 1 caggtgaactgcagcagtcaggacctgaactggtgagctgggtttcagtgaaagata 60
Db 79 CAGGTGAACATGCGAGCTGCGGCGAGAACTTGTAGGTGAGGACCTCAGTCAAGTTG 138
Qy 61 tcttgcagacttctgganacaaattcactgaatacacaccatgctcagctgggtgaagcagc 120
Db 139 TCCTGCACAGCTTCTGGCTTCAACATTAAGACTCCATATATGCACTGGTGTGAGGACGGG 198

Qy 121 catggaagagccttgagtgattgaggtatttaactcctaacaatggtgtaactaac 180
Db 199 CCTGAACAGAGCCCTGGAGTGGATTGGATGATGATCTCCTGAGAAATGGTATGATAATAT 258
Qy 181 aagcagaagtccaagggcagacattgactgtagacaagtcctccagcagcagctac 240
Db 259 GCCCGAAGTTCAGGGCAAGGCCACTTTTACTACAGACACATCCTCCCAACACAGCCTAC 318
Qy 241 atggagctccgcagcctgacatctgaggtattctgagctctattactgtgcaagagatact 300
Db 319 CTCAGCTCAGCAGCCTGACATCTGAGGACACTGCCGCTCTATTATTGTATAGGAGGACT 378
Qy 301 -----acgtccocgtttgttactgggtccaaaggagaccacggtcaccgtctctctca 351
Db 379 CCGACTGGGCCGTACTACTTTGACTACTGGGGCCAAAGGGACACCGGTACCGTCTCTCTCA 438
Qy 352 ggtggagcgggttcaggcggaggtggtctctggcgggtggcggtatcgagacatcgagctca 411
Db 439 GGTGGAGGCGGTTTCAGCGGAGGTGGCTCTGGCGGTGGCGGATCAGAAAAATGTGCTCA 498
Qy 412 cagtctccagcaatcatgtctgcatctccaggggagaaagtcaccatgacctgcaagtggc 471
Db 499 CAGTCTCCAGCAANTCATGTCTGCACTCTCCAGGGGAGAGGTCAACATAACTCGAGTGCC 538
Qy 472 agtcaagataagttacatgacatgacatgacatgacatgacatgacatgacatgacatgac 531
Db 559 AGTCAAGTGAAGTTACATGCACTGGTTCCAGCAAGAGGACGACACTTCTCCCAAACTC 618
Qy 532 tggattatgacacatccaaactggcttctgagtcctcctcgtcgtcagtgagtgagg 591
Db 619 TGGATTATAGACATCCAACTGGCTTCTGGAGTCCCTGCTCGCTTCAGTGCAGTGGA 678
Qy 592 tctgggacctcttattctcacaatcagcagcatgaggtgaggtgaggtgaggtgaggtgag 651
Db 679 TCTGGGACCTTACTCTCTCAATCAGCCGAATGGAGGCTGAAGATGCTGCCACTTAT 738
Qy 652 tactgcatcagcggagtagttaccctgcctcaactgctcgtcgtcgtcgtcgtcgtcgtcgt 711
Db 739 TACTGCCAGCAAGAGGAGTAGTTACCCACTCACGTTCCGTTGCTGGCAGTGGAGCTG 798
Qy 712 aaacgg 717
Db 799 AAACGG 804

RESULT 5

US-08-553-497A-19
; Sequence 19, Application US/08553497A
; Patent No. 5844093
; GENERAL INFORMATION:
; APPLICANT: KETLEBOROUGH, C. A.
; APPLICANT: BENDIG, MARY M.
; APPLICANT: ANSELL, KEITH H.
; APPLICANT: GUSSOW, DETLEF
; APPLICANT: ADAN, JAUME
; APPLICANT: MITJANS, FRANCESC
; APPLICANT: ROSELL, ELISABET
; APPLICANT: BLASCO, FRANCESC
; APPLICANT: PIULATS, JAUME
; TITLE OF INVENTION: ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLER, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD. SUITE 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible


```

: REFERENCE/DOCKET NUMBER: MERCK 1726
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-243-6333
: TELEFAX: 703-243-6410
:
: INFORMATION FOR SEQ ID NO: 21:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 732 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: cDNA
:
: HYPOTHETICAL: NO
:
: ANTI-SENSE: NO
:
: FRAGMENT TYPE: N-terminal
:
: ORIGINAL SOURCE:
: . ORGANISM: mouse
: STRAIN: Balb/c
: TISSUE TYPE: splenocytes
: IMMEDIATE SOURCE:
: CLONE: 3 D 3 (single-chain Fv, heavy chain)
: CLONE: plus linker)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..732
:
: US-08-553-497A-21

```

Query Match	70.2%	Score 503;	DB 3;	Length 732;
Best Local Similarity	83.6%;	Pred. No. 2.4e-133;		
Matches 611;	Conservative 0;	Mismatches 102;	Indels 18;	Gaps 3;
QY	2	aggtgaacctgcagcactcaggacctgaactgggtgnagcctggggcttcaatgaagatat	61	
Db	2	AGGTCCAACTGCGACGCTCAGGGCTGAACCTGGTGAAGCTGGGGCTTCAGTGAAGTTGT	61	
QY	62	cttgcgaagactcttgganacaaattcactgaatacacaccactcactgggtgaagcagagcc	121	
Db	62	CTGCAAGGCTCCGGCTACACCTTCACCGCCACTGGATGCATCGGTGAAGCAGAGGG	121	
QY	122	atggaaagacctgadtggattggaggtattaatcctaaacaattggtgtactaaactaca	181	
Db	122	CTGGACAAAGGCTTGAGTGGATCGAGAGGTTTAATCCGACGAACGGCCGTACTTAAC	181	
QY	182	acgagaagttcaagggcgaagggccacattgactgtagacaagtctccagcagaccctaca	241	
Db	182	ATGAGAAATATCAAGAGCAAGGCCACACTGACTGTAGACAAATCTCCAGCACAGCCTACA	241	
QY	242	tggagctccgcagcctgacatctgaggattcttcagctctattactgtgcaag-agatact	300	
Db	242	TGCAACTCAGCAGCCTGCATCTCAGGACCTCGCGGCTATTACTGTGCCAGTCGGGACT	301	
QY	301	acgggtcccg-----ttfcttactgggtccaaaggagaccaggtccacctctctct	349	
Db	302	ATGATTACGACGACGGTACTTTTGACTACTTGGGCCAAGGGACACCGGTACCGCTCTCT	361	
QY	350	cagggtgagcggttcagcgaggtggctctgcggtgcgcggtcggagatcggacatcagctca	409	
Db	362	CAGGTGCGGTGGCTCGGGCGGTGGTGGGTGGGTGGCGCGCGATCTGCATTGAGCTCA	421	
QY	410	ctcagttccagcaaatcatgtctgcattctccagggggagaaggtcaccatgacctgcagtg	469	
Db	422	CCCAGTCTCCAACAATCATGTCTGCATCTCCAGGGGAGAAGGTCAACCATGACCTGCAGTG	481	
QY	470	gcagctcaagataaagttacatgcactgggtaccagcagaagccctgtcacctcccccaaaa	529	
Db	482	ACAGCTCAAGTGAAGTTATCATGTACTTGGTACCGACGAGAAGACAGGATCTCCCGCCAGAC	541	
QY	530	gatggattatgacacatcaaaactggctcttgagttccctgcgtctcagtgagtgagtg	589	
Db	542	TCCTGATTTATGACACATCCAACTTGGCTTCTGGAGTCCCTGTTTCAGTGGCAGTG	601	
QY	590	gggtcgggacctcttatctctcacaatcagcagcatggagggctgtagatgctgccactt	649	

```

Db 602 GGCTGGGACCTCTTACTCTCTCACAATACGCCGAATGGAGGCTGAAGATGCTGCCACTT 661
Qy 650 attactccatcagcgagtagttacc-----gctcacgttcggtgctgggacacagt 703
Db 662 ATTACTGCCAGCATGGAGTAGTTACCCACCACCATGTATACATGTTTCGGAGGGGGACCAAGC 721
Qy 704 tggaaataaaa 714
Db 722 TGGAAATAAAA 732

RESULT 7
US-08-661-052-15
: Sequence 15, Application US/08661052
: Patent No. 5837243
: GENERAL INFORMATION:
: APPLICANT: Yashwant M. Deo
: APPLICANT: Joel Goldstein
: APPLICANT: Robert Graziano
: APPLICANT: Chezia Somasandaram
: TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
: TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street, Suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/661,052
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/484,172
: FILING DATE: 07-JUNE-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Arnold, Beth E.
: REGISTRATION NUMBER: 35,430
: REFERENCE/DOCKET NUMBER: MXI-043CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1679 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 11..1667
: US-08-661-052-15

```


TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS

Qy	422	caatcatgctctgcatctccaaggagggaagaaggtcaacatgacctgcagtgccagctcaagta	481
Db	413	CAATCATGTCTGATCTCCAGGGGAGAAGGTACCATGACCTGCAGTGGCCATGTCAAAGTG	472
Qy	482	taagtatactgactgataccagcagaagcctgtcaacctccccaaaaagatgattatg	541
Db	473	TAAAGGTACATGAAGTGGTTCACACAGAAGTCAGGCACCTCCCCAAAAGATGGATTATG	532
Qy	542	acacatccaaactggcttctggagtcctctgctgcttccagtgaggcagtgggctgggaacct	601
Db	533	ACACATCCAAACTGTCTTCTGGAGTCCCTGCTCGCTTCAGTGGCAGTGGGTCTGGGACCT	592
Qy	602	cttattctctcaatacacagatgagagcgtgaatctccaccttattactccatc	661
Db	593	CTTACTCTCTACAACTACGACGACATGGAGGCTGAAGATGCTGCCATTTATTACTGCCACG	652
Qy	662	agcggagtagttaccgcgtcacgttcgggtgctgggacacagttggaaaataaacg	716
Db	653	AGTGAGGTAGTATTCACCTCACCTTCGGTCTGGGACCAAGCTGGAGCTGAACG	707

RESULT 11

```

RESOL 11
US-08-190-199A-60
: Sequence 60, Application US/08190199A
: Patent No. 5830663
: GENERAL INFORMATION:
: APPLICANT: EMBLETON, Michael J.
: APPLICANT: GOROCHOV, Guy T.
: APPLICANT: JONES, Peter T.
: APPLICANT: WINTER, Gregory P.
: TITLE OF INVENTION: TREATMENT OF CELL POPULATIONS
: NUMBER OF SEQUENCES: 70
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
: STREET: 1100 New York Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005-3918
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Microsoft Word
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/190-199A
: FILING DATE: 13-JUL-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/GB92/01483
: FILING DATE: 10-AUG-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9212419.7
: FILING DATE: 11-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9117352.6
: FILING DATE: 10-AUG-1991
: INFORMATION FOR SEQ ID NO: 60:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 708 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..705
: US-08-190-199A-60.

```

Query Match	55.4%	Score	397;	DB	3;	Length	711;
Best Local Similarity	73.3%;	Prod. No.	2e-103;				
Matches	524;	Conservative	0;	Mismatches	182;	Indels	9;
Gaps	17;						
Y	2	aggtgaacctgcagcagctcagacactgaactggtgnagcctggggtctcagtgaaagatat	61				
Y	2	ATGTCAGCTGGTGTGAGTCTGGGGGAGGCTTAGTCGAGCCTGGAGGGTCCCGGAACACT	61				
D	62	ctctgaagactcttgganacaatctactgaatacaccatgcactcgggtgaagcagagacc	121				
D	62	CTGTGTCAGCCTCTGGATCTCACTTTCAGTAGCTTTGGAAATGCACATGGGTCTGTCAGGCTC	121				
Y	122	atggaaaagaccttgagttgaggtattaatcctaacaatagtggtactaactaca	181				
D	122	CAGAGAAGGGGTGGAGTGGGTGCGCATATATTAGTAGTGGCAGTAGTACCATCTACTATG	181				
Y	182	agcagaagtctcaaggcagggccacattgactgtgacaaagtctccagcacagcctaca	241				
D	182	CAGACACAGTGAAGGCCGATTCACCATCTCCAGAGACATCCCAAGACACCCCTGTTC	241				
Y	242	tggagctccgcagcctgcacattgaggtattctgcagttctattactgtgcagaagatacta	301				
D	242	TGCAAAATGACCAGTCTAAGGTCTGAGGACACGCGCATGTATTACTGTGCAAGAGATTAG	301				
Y	302	cgtgccggtttcttactaggttccaaggacacacaggtcacccgtctcttcaggtgagcgc	361				
D	302	GG-----GCTATTATGGGGCCAAAGGACACTGTGCTACTGTCTCTCGAGGTGTGGTG	352				
Y	362	gttcaggcggaggtggctctggcttggcggtatcggaatcggaatcgagctcactcagttccag	421				
Y	353	GTAGCGGTGGTGGCGCACTGGCGGGGGCTCTCAAAATGTTCTTCAACCACTCTCAG	412				

Query Match 50.3%; Score 360.8; DB 4; Length 756;
Best Local Similarity 70.5%; Pred. No. 3.5e-93;
Matches 520; Conservative 0; Mismatches 194; Indels 24; Gaps 2;

INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: PW215-51
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1911 /note= "64-87 FLAG peptide and
OTHER INFORMATION: enterokinase cleavage site;97-456 FWP51 heavy
OTHER INFORMATION: chain variable domain;457-501 15 aa linker
OTHER INFORMATION: sequence;502-822 FWP51 light chain variable
OTHER INFORMATION: domain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 1..63
OTHER INFORMATION: /note= "ompA signal peptide"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 94..1911
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1912..2012
OTHER INFORMATION: /function= "3' non-coding region of
OTHER INFORMATION: the exotoxin A gene"
FEATURE:
NAME/KEY: misc feature
LOCATION: 826..1911
OTHER INFORMATION: /note= "Exotoxin A gene coding
OTHER INFORMATION: region (coding for amino acids 252 to 613 of the
OTHER INFORMATION: mature exotoxin A)"
US-08-235-838-15

Query Match 48.6%; Score 348.8; DB 1; Length 2012;
Best Local Similarity 70.5%; Pred. No. 1.2e-89;
Matches 512; Conservative 0; Mismatches 199; Indels 15; Gaps 3;
QY 1 caggtgaactcagcagctcagcagcctgaactggtgnagcctgggcttcagtgaaagata 60
DB 97 CAGGTACAACTGCAAGCTCTGGGCTGAGCTGGTGGGCTTGGGACTTCAGTGAAGCTG 156
QY 61 tcttgaagactcttgganacaaattcactgaatacacaccatgctggtggaagcagagc 120
DB 157 TCCTGCAAGCTTCTGATTACACCTTCACCAGCTACTGGATGAAGTGGTGAAGCAGG 216
QY 121 catggaagagccttgagtgagtgattgagtgatttaactcctaacaatggtggtactaac 180
DB 217 CCTGGCAAGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 276
DB 181 aagcagaagttcaaggcaggaagccacattgactgtagacaagtcctccagcacagcctac 240
DB 277 AATCAAAATGTTCAAGGACAAAGGCGCATGACTGTAGACAAGTCTCCAAATACAGCCTAC 336
QY 241 atgagctccgcagcctgacatctgagctgagctgagctgagctgagctgagctgagctg 300
DB 337 ATGCAACTCAGCAGCTGACATCTGAGACTCTGCGGCTCTTACTGTGCAAAAGGGGGG 396
QY 301 acggtccgc-----ttgttactgggtccaaaggacacgggtcacccgtctctctca 351
DB 397 GCCTCTGGGACTGGTACTTCGATGCTGGGGCCCAAGGACACAGGTCACCGTTTCCTCT 456
QY 352 ggtggaagcgttcaggcggaggtggtctggtgggtggcgatcgagacatcgagctcaact 411
DB 457 GCGGCTGGCGGTTCTGTGTGGGTGGTTCGCGGGGTGGCGGTTCTGACATCCAGCTGACC 516
QY 412 cagtcctccagcaatcatgtctgcatctccagggagagaggtcaccatgacgtcgagtggc 471
DB 517 CAGTCTCCATCCCTCAGTCTGTCATCTCTGGGAGGCCAAGTCAACATCACTTGCAGGCA 576
QY 472 agctca---agtataagttacatgactggttacacagcagaagcctgtcacctccccaaa 528
DB 577 AGCCAAGACATTAAGAAGTATATAGTTGTTACCAACACAAGCCTGGAAAAAGTCCTCGG 636
QY 529 agatggatttatgacacatccaaactggtcttctggagtcctcctgctcagtggcagt 588

DB 637 CTACTCATACACTACACATCTGTATTACAGCCAGGATCCCATCCAGGTTTCAGTGGAACT 896
QY 589 gggctctgggagagattattcttctcacaatcagcagcagcagcagcagcagcagcagc 648
DB 697 GGGTCTGGGAGAGATTATTCTTCTCAGCATCCCAACCTGGAGCTGAAGATATTGCAACT 756
QY 649 tattactgccaatcagcggagtagttaccacgctcacgcttcgggtgctgggacacagttggaa 708
DB 757 TATTATG---TCTACATTATGATTATCTGTACACGTTCCGGAGGGGGCCACCAAGCTGGAG 813
QY 709 ataaaa 714
DB 814 ATCAAA 819

Search completed: November 30, 1999, 14:10:18
Job time: 5144 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 1999, 12:44:33 ; Search time 545.98 Seconds
(without alignments)
5615.269 Million cell updates/sec

Title: US-08-940-544-3

Perfect score: 717

Sequence: 1 caggtagaactgcagcagtc.....cacagttggaataaaacgg 717

Scoring table: IDENTITY_NUC

Searched: 780561 seqs, 2137953050 residues

base :

GenEmbl.:

1: gb_bal.*
2: gb_ba2.*
3: gb_cm.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pil.*
8: gb_pil2.*
9: gb_pri.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_sts.*
14: gb_sy.*
15: gb_un.*
16: gb_vi.*
17: em_fun.*
18: em_htg.*
19: em_hum1.*
20: em_hum2.*
21: em_in.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_sy.*
30: em_un.*
31: em_vi.*
32: gb_htg1.*
33: gb_htg2.*
34: gb_in1.*
35: gb_in2.*
36: gb_in3.*
37: gb_in4.*
38: gb_in5.*
39: gb_in6.*
40: gb_in7.*
41: gb_in8.*
42: gb_in9.*
43: gb_in10.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query
No. Score Match Length DB ID Description

1	590.4	82.3	723	12	AF035617	AF035617 Mus muscu
2	522.2	72.8	726	5	A45598	A45598 Sequence 29
3	521.6	72.7	726	5	A45594	A45594 Sequence 25
4	521.6	72.7	726	5	AR063195	AR063195 Sequence
5	520.4	72.6	732	5	E13410	E13410 cDNA encodi
6	517.8	72.2	2364	5	E13412	E13412 cDNA encodi
7	517.4	72.2	1797	5	I84705	I84705 Sequence 2
8	512.6	71.5	726	5	A45596	A45596 Sequence 27
9	512.6	71.5	726	5	AR063196	AR063196 Sequence
10	511.8	71.4	717	12	AF025535	AF025535 Mus muscu
11	508	70.9	723	14	AF169027	AF169027 Synthetic
12	507.2	70.7	810	5	A45006	A45006 Sequence 1
13	505.2	70.5	726	12	MMU20617	U20617 Mus musculu
14	504.6	70.4	732	5	A45588	A45588 Sequence 19
15	504.6	70.4	732	5	A45600	A45600 Sequence 31
16	504.6	70.4	732	5	AR063192	AR063192 Sequence
17	503	70.2	732	5	A45590	A45590 Sequence 21
18	503	70.2	732	5	AR063193	AR063193 Sequence
19	502.8	70.1	738	12	AF074900	AF074900 Mus muscu
20	501.6	70.0	1679	5	AR054409	AR054409 Sequence
21	493.6	68.8	798	12	AF000955	AF000955 Mus muscu
22	487	67.9	738	5	A45592	A45592 Sequence 23
23	487	67.9	738	5	AR063194	AR063194 Sequence
24	456.6	63.7	1944	14	XXU78761	U78761 Synthetic c
25	449.2	62.6	678	12	AF003714	AF003714 Mus muscu
26	445.4	62.1	596	12	AF003724	AF003724 Mus muscu
27	438	61.1	684	12	AF003720	AF003720 Mus muscu
28	434.8	60.6	696	12	AF003721	AF003721 Mus muscu
29	429.6	59.9	687	12	AF003725	AF003725 Mus muscu
30	421.6	58.8	702	12	AF003708	AF003708 Mus muscu
31	418.6	58.4	738	12	AF141321	AF141321 Mus muscu
32	414.2	57.8	729	12	AF002242	AF002242 Mus muscu
33	413.2	57.6	738	12	MMU40581	U40581 Mus musculu
34	412.4	57.5	720	12	MMU40582	U40582 Mus musculu
35	412.4	57.5	4354	14	ASY14583	Y14583 Artificial
36	410.8	57.3	738	5	A41674	A41674 Sequence 1
37	403.4	56.3	684	12	AF003711	AF003711 Mus muscu
38	397	55.4	711	5	A25748	A25748 Variable re
39	397	55.4	711	5	AR051416	AR051416 Sequence
40	391.4	54.6	783	12	AB007986	AB007986 Mus muscu
41	390.6	54.5	1314	5	A68604	A68604 Sequence 4
42	390	54.4	684	12	AF003716	AF003716 Mus muscu
43	388.6	54.2	783	12	MMU010065	MMU010065 Mus muscu
44	386.8	53.9	708	5	A25744	A25744 Variable re
45	386.8	53.9	708	5	AR051414	AR051414 Sequence

ALIGNMENTS

RESULT 1

AF035617
LOCUS AF035617 723 bp mRNA ROD 30-SEP-1999
DEFINITION Mus musculus scFv antibody SS mRNA, partial cds.
ACCESSION AF035617
NID 92921393
VERSION AF035617.1 GI:2921393
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 723)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Chowdhury, P.S., Viner, J.L., Beers, R. and Pastan, I.

TITLE Isolation of a high-affinity stable single-chain Fv specific for mesothelin from DNA-immunized mice by phage display and construction of a recombinant immunotoxin with anti-tumor activity
Proc. Natl. Acad. Sci. U.S.A. 95 (2), 669-674 (1998)

JOURNAL 98118570

MEDLINE 2 (bases 1 to 723)

REFERENCE Chowdhury, P.S.

AUTHORS Direct Submission

TITLE Submitted (24-NOV-1997) LMB, NCI, 37 Convent Drive, Bld. 37 Rm.

Result Query
No. Score Match Length DB ID Description


```

QY 242 tggagctccgagcctgacatctgaggtattctgagttctattactgtgcaag-agataact 300
Db 242 TCGAACTCAGCAGCGCTGACATCTAGGACTTCGCGTCTATTACTGTGCCAGTCGGGACT 301
QY 301 acgggtcccg-----tttgcttactgggtccaaggagaccaggtccacgtctcct 349
Db 302 ATGATTACGACGGACGGTACTTTGACTACTGGGGCCAGGGGACACACGGTACCCTCTCT 361
QY 350 caggtggagggcgttcaggcgaggtggtctggtggcggtggcggtgagatcgacatcgagctca 409
Db 362 CAGGTGGCGGTGGCTCGGGCGGTGGTGGTGGCGGGTGGCGGGATCTGACATTTAGCTCA 421
QY 410 ctcaattcccaacaatctgtctgcatctccaaggaggagaaggtcaaccatgaactcagtg 469
Db 422 CCGAGTCTCCATCAATCAATGCTGATCTCCAGGGGAGAGGTCAACATGACCTCAGTG 481
QY 470 cgagctcaagtaagttacatgctggtaccagcagaagcctgtcacctcccccacaaa 529
Db 482 ACAGCTCAAGTGTAGTTACATGTACTGGTACCAGACAGACAGATCCCTCCCCAGAC 541
QY 530 gatgatttatgacacatccaaactgggtcttggagtcctctgctgcttcagtgagtg 589
Db 542 TCCTGATTTATGACACATCCAACTGGCTTCGGAGTCCCTGTTTCAGTGGCAGTG 601
QY 590 ggtctggagaccttattctctcaaatcagcagcatgagagcgttagatgctgcaactt 649
Db 602 GGTCTGGGACCTTTACTCTCTCAATCAGCCGAATGGAGCTGAAGATGTGCCACTT 661
QY 650 attactgcatacagcaggaagtaactaccgctcacttgcgttcggtcgggacagattggaaa 709
Db 662 ATTACTGCAGCAGTGGAGTAGTTACCCACACACACCTTCGGTCTGGGACCAAGCTGGAAA 721
QY 710 taaaa 714
Db 722 TAAAA 726

RESULT 3
LOCUS A46594 726 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 25 from Patent WO9525167.
ACCESSION A46594
NID 92300759
VERSION A46594.1 GI:2300759
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 726)
AUTHORS Kettleborough,A.C., Bendig,M.M., Ansell,K.H., Guessow,D., Adan,J.,
Mitjans,F., Rosell,E., Blasco,F. and Pluats,J.
TITLE ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR ANTIBODIES
JOURNAL Patent: WO 9525167-A 25 21-SEP-1995;
MERCK PATENT GMBH (DE)
COMMENT Other publication ZA 9502174 951227
Other publication CN 1124501 960612
Other publication HU 73461 960828
Other publication CZ 9503014 960214
Other publication PL 311661 960304
Other publication CA 2163012 950921
Other publication NO 954626 951116
Other publication AU 2071695 951003.
Location/Qualifiers
1..726
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/tissue_type="SPLENOCYTES"
1..726
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"

CDS

```

```

/codon_start=1
/protein_id="CAA02930.1"
/db_xref="PID:e306173"
/db_xref="PID:g2300760"
/db_xref="GI:2300760"
/translation="OVKLOESGAELVKPGASVLSKASGYTFTSHMMHWKORAGG
LEWTEINPRTAPNTNNEKPKSKATLTVDKSSSTAYWQLSSLSESAVYXCASRDYD
YDRYDYWGCGTIVTVSSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
DSSSVSYWYQKTSPPRLLIYDTSNLASGVPRFSGSGSGTYSLSLTISRMEADA
ATYCCQWSSYPLTFGATKLEIK"
BASE COUNT 174 a 192 c 206 g 154 t
ORIGIN

```

```

Query Match 72.7%; Score 521.6; DB 5; Length 726;
Best Local Similarity 84.4%; Pred. No. 1.4e-150;
Matches 613; Conservative 0; Mismatches 101; Indels 12; Gaps 2;

QY 1 caggtgaaactgcagcagctcaggaactggtgagcctggggtcagtggaagata 60
Db 1 CAGGTGAAACTGCAAGGAGTCTGGGGCTGAAGTGTGAAGCCTGGGGCTTCAGTGAAGTTG 60
QY 61 tctgtcaagactctgaaacaaatttcactgaatacacaccatgcactgggtgaagcagagc 120
Db 61 TCTGTCAAGGCTTCGGGCTACACCTTCACAGCCACTGGATGCACCTGGGTGAACAGG 120
QY 121 catggaagagccttgagtgagtgattggaggtatttaacttaacaaatggtggtactaaactac 180
Db 121 GCTGCACAAGCCCTTGAAGTGGATCGGAGATTAATCCAGAACGGCGCTTACTTAACACTAC 180
QY 181 aagcagaagttcaaggcgaaggccacattgactgtagaagtcctccagcacagcctac 240
Db 181 AATGAGAAATTCAGAGACAGAGCCACACTGACTGTAGACAAATCCTCAGCACAGCCCTAC 240
QY 241 atggagctccgcagcctgacatctgaggtattctcagctctattactgtgaag-agatac 299
Db 241 ATGCAACTCAGCAGCCTGACATCTGAGGACTCTCGGGTCTATTACTGTGCGAGTCGGGAC 300
QY 300 tacgggtcccg-----tttgcttactgggtggtccaaaggagaccaggtcacggtctcc 348
Db 301 TATGATTAGCAGCGCAGGTACTTTGACTACTGTGGGCCAAGGACACAGGTCACCGTCTCC 360
QY 349 tcaggtgagcggttcaggcgagggtggctctgctggcggtggcggtcggagatcgagcactc 408
Db 361 TCAGGTGGCGGTGGCTCGGGCGGTGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 420
QY 409 actcagttccagcaatcatgtctgcattccaggggagaggtcaccatgacctgcagt 468
Db 421 ACCCAGTCTCCAAACAATCATGTCTGCAATCTCCAGGGAGAGAGGTACCATGACCTGCAGT 480
QY 469 ggcagctcaagataaagttaacatgctggttaccagcagaagcctgtcacctcccccaaa 528
Db 481 GACAGCTCAAGTGTAGTTACACCTACTGTTACAGCAGAGACAGGATCCTCCCCCAGA 540
QY 529 agatggatttatgacacatccaaactggcttctctggagtcctcgtcgtcgttcagtgagtg 588
Db 541 CTCCTGATTTTATGACATCCAACTGGCTTCTGGAGTCCCTGTTTCGCTTCAGTGGCAGT 600
QY 589 gggctctgggaacctttattctctcaaatcagcagcatggaggtgtagatgtgccact 648
Db 601 GGGCTGGGACCTTCTTACTCTCTCACAATCAGCCGAATGGAGGCTGAAGATGTGCCACT 660
QY 649 tattactgccaatcagcgagtagttaccgcgtcagttcgggtgctgggacacagttggaa 708
Db 661 TATTACTGCCAGCAGTGGAGTAGTTACCCGCTCAGTTCGTTGGTGGTGGTGGTGGTGGTGG 720
QY 709 ataaaa 714
Db 721 ATAAAA 726

RESULT 4
AR063195

```

QY	709 ataaaa 714'
Db	721 ATAAAA 726
RESULT	5
EL3410	
LOCUS	732 bp DNA PAT 24-JUN-1998
DEFINITION	cDNA encoding single chain antibody from mouse which have binding ability with human growth hormone (HGH).
ACCESSION	E13410
NID	9352215
VERSION	E13410.1 GI:3252215
KEYWORDS	JP 1997187281-A/1.
SOURCE	Mus musculus.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS	Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	1 (bases 1 to 732)
JOURNAL	Masuda, S., Tatsumi, H. and Koyama, T. ANTIBODY-FREELY LUCIFERASE FUSION PROTEIN, GENE OBTAINED THEREFROM, NOVEL RECOMBINANT DNA AND PRODUCTION OF ANTIBODY-FIREFLY LUCIFERASE FUSION PROTEIN Patent: JP 1997187281-A 1 22-JUL-1997; KIKKOMAN CORP
COMMENT	OS Mus musculus (mouse) PN JP 1997187281-A/1 PD 22-JUL-1997 PF 09-JAN-1996 JP 1996001812 PI MASUDA SUSUMU, TATSUMI HIROKI, KOYAMA TAIJI PC C12N13/09,A61K39/395,C07H21/04,C07K14/47,C07K19/00,C12P21/02, PC GOIN33/535// PC C12N9/02,(C12P21/02,C12R1/19); CC strandedness: Double; CC topology: Unknown; CC hypothetical: No; CC anti-sense: No; FH Key Location/Qualifiers
FT	source 1..732 /organism='Mus musculus' /clone='pSCFV4'
FT	mat_peptide 1..732 /product='Single chain antibody from mouse FT which have binding ability with human growth hormone FT (HGH)'
FT	misc_feature 1..354 /product='H chain variable region of single chain antibody from mouse'
FT	misc_feature 355..399 /product='linker peptide'
FT	misc_feature 400..726 /product='L chain variable region of single chain antibody from mouse.'
FEATURES	Location/Qualifiers 1..732 /organism="Mus musculus" /db_xref="taxon:10090"
BASE COUNT	179 a 183 c 212 g 158 t
ORIGIN	
Query Match	72.6%; Score 520.4; DB 5; Length 732;
Best Local Similarity	84.5%; Pred. No. 3.2e-150;
Matches 610:	Conservative 0; Mismatches 403; Indels 9; Gaps 2;
Qy	2 aggtgaactgcagcagctcgaggacctgaactggtgtagccctggggcttcagtgaagatat 61
Db	8 AGGTGAAGCTGCAGGAGCTCAGGAGCTAGCTAGCTGGTGGAGCGCTTGCTCTCAGTGAAGATT 67


```
Qy 536 ttatgacacatcaaacctgcttctgagtcgctccctgctcgttcagtgagcagtggtctg 595
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 545 TTTATGGCACATCCAACTGGCTTCTGGAATCCCTGTTTCGTTTCAGTGGCAGTGGATCTG 604
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 596 ggacctctattcttcacatcaagcagcatggagctgtagatgctgctccacttattact 655
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 605 GGACCTCTATTCTCTACATCAGCACATGGAGGGTGAAGATGCTGCCACTTATTACT 664
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 656 gccatcagcggagtagttaccgcctcacgttcctgctgctggtggacacagtggaataaac 715
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 665 GCCAGCAGTGGAGTAGTAACCATTCACGTTCTCGGCTCGGGGACCAAGCTGGAATAAAG 724
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 716 g 716
Db 725 G 725

RESULT 7
US 184705 184705 1797 bp DNA PAT 04-APR-1998
DEFINITION Sequence 2 from patent US 5696237.
ACCESSION I84705
NID 93022225
VERSION I84705.1 GI:30222225
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1797)
AUTHORS Fitzgerald, D., Chaudhary, V. Kumar, Pastan, I. Harry,
Waldmann, T. Alexander and Queen, C.L.
TITLE Recombinant antibody-toxin fusion protein
JOURNAL Patent: US 5696237-A 2 09-DEC-1997;
FEATURES
Location/Qualifiers
1..1797
/organism="unknown"
BASE COUNT 336 a 598 c 571 g 292 t
ORIGIN

Query Match 72.2%; Score 517.4; DB 5; Length 1797;
Best Local Similarity 83.7%; Pred. No. 3.1e-149;
Matches 597; Conservative 0; Mismatches 113; Indels 3; Gaps 1;

Qy 1 caggtgaactgcagcagtcaggacctgaactggtgagcctgggttcagtgaaagata 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 4 CAGGTCCAGCTGCAGCAGCTCTGGGCTGAACCTGGCAAAACCTGGGGCTCAGTCAAGATG 63
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
61 tccctgaagacttctgganacaaattcactgaataacacccatgcactgggtgaagcagagc 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 64 TCCTGCAAGGCTTCTGGCTACACCTTTACTAGCTACAGGATGCAGTGGTAAACAGAGG 123
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 121 catggaagagccttgatgagtgaggtattatctcctaacaatggtggtactaactac 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 124 CCTGGACAGGCTCGGAATGATGGATATATTAACTCAGCAGCTGGTATCTACTCAATAC 183
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 181 aagcagaagttcaaggcgaagccacattgactgtagaagtcctccagcagacgctac 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 184 AATCAGAAGTTCAAGGACAGGACAGCCACATTGACTGCAGACAAATCCTCCAGCACAGCCTAC 243
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 241 atgagctcccgagcctgacatctgaggtattctcagcttattactgtgaagagatact 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 244 ATGCAACTGACGACCTGACATTTGAGGACTCTGCAGTCTATTACTGTGCAAGAGGGGGG 303
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 301 acggtccctgttctgactggttcacaggaccacggtccacgctctcctcaggtggaggc 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 304 GGGGTC---TTTGACTACTGGGGCCAGGAACCACTCTCAGACTCTCTCCGGAGGGCGGT 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 361 ggttcagcggaggtgctctggtgggtggatcgagacatcgagctcactcagctctcca 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 361 GGGTCGGGCGGTGGGGCTCGGGTGGGGCGGCTCTCAAAATTGTTCTCACCCAGCTCTCCA 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 421 gcaatcagctgcatctccaggggaggaaggtccaccatgacctcagtgagcagctcaagt 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

```
Db 421 GCAATCATGTCATCTCCAGGGGAGAGGTCAACATAACCTGCAGTGCAGCTCAAGT 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 481 ataagttacatgcactgtaccagcagaaagcctgtcacctcccccaaaagatgattat 540
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 481 ATAAGTTACATGCATGCTGTTCCAGCAGAACCCAGGCACTTCTCCCAAACTCTGGATTAT 540
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 541 gacacatcccaaaatggcttctgagtcctcctgctcgttcagtgagtggtggtggacc 600
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 541 ACCACATCCAACTGGCTTCTGGAGTCCCTGCTGCTTCAGTGCAGTGCAGTGGGACC 600
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 601 tcttattctctcaaatcagcagcatggagctgtagatgctgctccacttattactgcat 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 601 TCTTACTCTCTCAATCAGCCGAATGGAGGCTGAAGATGCTGCCACTTATTACTGCGCAT 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 661 cagcggagtagttaccgcctcacgttcggtggtggacacacagtggaataaaa 713
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 661 CAAGGAGTACTTACCACCTACGTTCTGGTCTGGGACCAAGCTGGAGCTCAA 713
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 8
A46596 726 bp DNA PAT 07-MAR-1997
LOCUS Sequence 27 from Patent WO9525167.
DEFINITION A46596
ACCESSION 92300761
NID A46596.1 GI:2300761
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 726)
AUTHORS Rodentia; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Kettleborough, A.C., Bendig, M.M., Ansell, K.H., Guessow, D., Adan, J.,
Mitjans, F., Rosell, E., Blasco, F. and Piuels, J.
TITLE ANTI-EGFR SINGLE-CHAIN FVS AND ANTI-EGFR ANTIBODIES
JOURNAL Patent: WO 9525167-A 27 21-SEP-1995;
COMMENT Other publication ZA 9502174 951227
Other publication CN 1124501 960612
Other publication HU 73461 960828
Other publication CZ 9503014 960214
Other publication PL 311661 960304
Other publication NO 2163012 950921
Other publication NO 954626 951116
Other publication AU 2071695 951003.
FEATURES
Location/Qualifiers
1..726
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/tissue_type="SPLENOCYTES"
/clone_lib="7 G 1 (SINGLE-CHAIN FV, HEAVY, LIGHT CHAIN,"
/clone_lib="LINKER)"
1..726
/notes="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA02931.1"
/db_xref="PID:8306174"
/db_xref="PID:92300762"
/db_xref="GI:2300762"
CDS
/translation="EVKIQSGAELVKPGASVKLSCKASGYTFTSHLDHWKQGWOG
LEWIGQFNPNNGRTNNEKFKSKATLTVDKSSSTAYIELSSLTSDCSVYVCASRDYD
YGRYFDYWGQCTVTTVSSGGSGGGGGGGSDIELTQSPITMSAPGKVTWICS
DSSSYMYWYQOKTCSRPRLIYDTSNLASGVPVRFSGSGSGTSTYSLTISRMEADA
ATYYCQWSSYPLTFGAGTKLEIK"
BASE COUNT 171 a 193 c 205 g 157 t
ORIGIN

Query Match 71.5%; Score 512.6; DB 5; Length 726;
Best Local Similarity 83.7%; Pred. No. 8.2e-148;
```


Db	421	CCAGCAATCATGTCTGCATCTCCAGGGGGAAGGTCACCAATAACCTGCAGTGCAGCTCA	480
Qy	478	agtataagttacatgcactggtaccagcagaagccctgtcaacctccccaaaaagatggatt	537
Db	481	AGTTAAGTTTACATGCACACTGGTTCCACGAGAAGCCAGGCACCTTCTCCCAAACCTCGGATT	540
Qy	538	tatgacacatccaactggcttcttgaggctccctgctcgcttcagtcagtcagtcgggtctggg	597
Db	541	TATACCACATCCAACTGGCTTCTTGAGTCCCTGCTCGCTTCAGTGGCAGTGGATCTGGG	600
Qy	598	accttatttcttcacaatcacagcatgatgagcgctgtagatgctgcaccttattctgc	657
Db	601	ACCTCTTACTCTCTCAACAATCACGCCAATGAGCGCTGAAGATGCTGCGACTTATTACTGC	660
Qy	658	catacgcgaagttagttaccgcgtcacgcttcgctcgctgggacacagcttggaataaaa	714
Db	661	CAGCAAGAGTAGTATTACCCCGCTACAGTTCGGTCTGGGACCAAGCTGGAGCTGAAA	717


```

RESULT 11
AF169027      AF169027      723 bp      mRNA      SYN      25-AUG-1999
LOCUS        Synthetic construct anti-CEA single chain monoclonal antibody C50
DEFINITION
ACCESSION    AF169027
NID          95764462
VERSION      AF169027.1 GI:5764462
KEYWORDS     synthetic construct.
SOURCE       synthetic construct.
ORGANISM     artificial sequence.
REFERENCE   1 (bases 1 to 723)
AUTHORS     Zhang,M. and Kong,J.
TITLE       Construction and screening of the anti-CEA single chain antibody
            gene
JOURNAL     Shengwu Huaxue Yu Shengwu Wuli Jinzhan 23 (5), 470-474 (1996)
REFERENCE   2 (bases 1 to 723)
AUTHORS     Zhang,M. and Kong,J.
TITLE       Direct Submission
JOURNAL     Submitted (15-JUL-1999) Department of Immunodiagnosics, National
            Vaccine and Serum Institute, Chaoyang District, Beijing, Beijing
            100024, China
FEATURES
Source       Location/Qualifiers
             1..723
             /organism="synthetic construct"
             /db_xref="taxon:32830"
             /note="obtained from Mus musculus strain BALB/c hybridoma"
misc_feature 1..360
             /note="Kappa light chain variable region"
             <1..>723
             /codon_start=1
             /transl_table=11
             /product="anti-CEA single chain monoclonal antibody C50"
             /protein_id="AAD51317.1"
             /db_xref="PID:G5764463"
             /db_xref="GI:5764463"
             /translation="EVHLOQSALVELRGSVKLSCTASGFNIKHYMHVKKORPQG
LEWIGWPENVDTEAPRKATMTADTSNTAYLQLSLTSEDYAVYYCNHYRYA
VGGLDLYGQGTIVTVSGGGGGGGGGSDIELTGSPAIMSASPKGVMTCSA
SSSVSYIHYYQQKSGTSPKRWYDTSKLASGVPARFSGSGSTYSLTISTMEAEVA
TYICOQNPNPYTFGGGKLEIK"
misc_feature 91..105
             /note="encodes CDR1"
misc_feature 148..198
             /note="encodes CDR2"
misc_feature 295..327
             /note="encodes CDR3"
misc_feature 361..405
             /note="artificial linker"
misc_feature 406..721
             /note="heavy chain variable region"
misc_feature 475..504

```

```

misc_feature      /note="encodes CDR1"
550..570
misc_feature      /note="encodes CDR2"
570..593
misc_feature      /note="encodes CDR3"
BASE COUNT      179 a 189 c 195 g 160 t
ORIGIN
Query Match      70.9%   Score 508;   DB 14;   Length 723;
Best Local Similarity 82.5%;   Pred. No. 2.2e-146;
Matches 596; Conservative 0; Mismatches 117; Indels 9; Gaps 1;

```

QY	2	aggtgaactgcagcagtcaggaccctgaacctgctgnaagctctgggggttcctcaagtgaagatat	61
Db			
Db	2	AGGTTACCTTCGACAGACTCTTTGGCAGAGGCTTGAGAGTTCAGGGGCTTCAGTCAAGTTGT	61
QY	62	cctgcaagactctctganacaaattccactgaatacacaccatgcactgggtgaagcagacc	121
	62	CTGTGCACAGCTTCTGGCTTCAACATTAACACACTACTATATGCACTGGGTGAACACAGAGC	121
QY	122	atggaagaagccttgagtgagtgagggtatttaactctaacaatggtggtactaactaca	181
Db			
Db	122	CTGAACAGGCCCTGGAGTGGATGGATTAATCCTTGAGATGTTGATACGTGATG	181
QY	182	agcagaagttcaagggoaaggccacattgactgtagacaagtcctccacacagcctaca	241
Db			
Db	182	CCGCCAAGTTCACAGGGCAAGGCCACTATGACTCGAGACACATCCTCCAACACAGCGCTACC	241
QY	242	tggagctccgcagcctgacatctcgagattctcgagttctattactgtgcaagagata---	299
Db			
Db	242	TGCAGCTCAGCAGCCTGCATCTCGAGGACACTCCGCTCTATTACTGTAATCACTATAGGT	301
QY	299	-----ctacggtcccggttctgacttgggtccaaggagccacggtcaccgtctcctcag	352
Db			
Db	302	ACGCCGTAGGGGGTGTCTTGGACTACTGGGGTCAAGGCACCGGTCAACCGTCTCCTCAG	361
QY	353	gtggaggcggttcaggcgagggtgggtctbgtgggtggcggtacgagacatcgagctcactc	412
Db			
Db	362	GTGGAGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCGGACATCGAGCTCAGCTC	421
QY	413	agttccagacaatcatgtctgcattccaggggaggaaggtcaccatgcacctgcagtgcca	472
Db			
Db	422	AGTCTCAGCAATCATGCTGTCATCTCCAGGGAGAAGGTCCACATGACTCGAGTGCCA	481
QY	473	gctcaagataaagttacatgcactggtaccagcagaagcgtgtcactctcccccaaaagt	532
	482	GCTCAAGTGTAAATTACATACACTGGTATCAGCAGAGTTCAGGCACCTCCCCCAAAAGAT	541
QY	533	ggattatgacacatccaaactggctctggagtcctgctcgcttcagtgagctgggt	592
Db			
Db	542	GGGTTTATGACACATCCAACTGGCTTCGTGGAGTCCCTGCTCGTTCAGTGGCAGTGGGT	601
QY	593	ctgggacctttattctcacaaatcagcagcatggagcgttagatgctgcaccttatt	652
Db			
Db	602	CTGGGACCTCTTACTCTTCACAATCAGCACCAATGAGAGGCTGAAGTAGTGGCACTTATT	661
QY	653	actgcatcagcgaggtagttaccgctcacgttcggtcggtggtggtggtggtggtggtggt	712
Db			
Db	662	ACTGCCACGAGTGGATATAAACCCATACACGTTCCGGAGGAGGACCAAGCTGGAAATAA	721
QY	713	aa	714
Db			
Db	722	AA	723

RESULT	12
A45006	A45006
LOCUS	810 bp DNA
DEFINITION	Sequence 1 from Patent WO9515341.
ACCESSION	A45006
NID	g2299557
VERSION	A45006.1 GI:2299557
PAT	07-MAR-1997

```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 810)
Chester,R.A., Hawkins,R.E. and Begent,R.H.
ANTIBODY AGAINST CARCINOEMBRYONIC ANTIGEN (CEA)
Patent: WO 9515341-A 1 08-JUN-1995;
CANCER RES CAMPAIGN TECH (GB)
Other publication CA 2177584 950608
Other publication AU 1194795 950619.

FEATURES
source
1..810
/organism="unidentified"
/db_xref="taxon:32644"

CDS
1..810
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA02813.1"
/db_xref="PID:e306097"
/db_xref="PID:g2299568"
/db_xref="GI:2299568"
/translation="EVVINKYLLPTAAAGLLLLAOPAMAQVKLOQSGAELVRSGTSS
KLSCTASGNIKDYSVHMLWRQGLEWIGTDPNGDTYAPKFGQKATFTTIDTSS
NTAYLQLSLTSEDATVYCNGETPTGYPYFDYWGOGTTVTYSSGGGGGGGGGGGG
SENVLQSPALMSASGEKVTITCSASSVSVMHWFQKPGTSPKLIWIIYSTSLWASG
PARFSGSGGTSTVSLFISRMEDADAAITYCQQRSSYPYLTFTGAGTKLELRAA"

BASE COUNT      189 a      220 c      222 g      179 t
ORIGIN

Query Match          70.7%; Score 507.2; DB 5; Length 810;
Best Local Similarity 82.2%; Pred. No. 3.9e-146;
Matches 597; Conservative 0; Mismatches 120; Indels 9; Gaps 1;

```

Query Match	70.7%	Score 507.2	DB 5	Length 810
Best Local Similarity	82.2%	Pred. No. 3.9e-146		
Matches 597	Conservative	0	Mismatches 120	Indels 9
Gaps	1			
QY	1	caggtgaacctgcagcagctcagacacctgaactggtgagcctgggcttcagtgaaata	60	
Db				
Db	79	CAGGTGAACCTGACAGCTCTGGGCAGAACTGTGAGGTCAGGGACCTCAGTCAAGTTG	138	
QY	61	tctgtcaagactcttgganacaaattcactgtaatacaccatgcactgggtgaagcagagc	120	
Db				
Db	139	TCCTGTCACAGCTTCTGGCTTCAACATTAAAGACTCCTATATGCACCTGGTTGAGCGAGGG	198	
QY	121	catggaagaagccttaagtgagattggaagttaataactcaactaactggtgactaaactac	180	
Db				
Db	199	CCTGAACAGGCGCTGGAGTGGATGGATGGATTGATCCTGAGAAATGGTGATCTGAATAT	258	
QY	181	aagcagaagttcaaggscgaagggccacattgactgtagacaaagtctccacgacacgctac	240	
Db				
Db	259	GCCCCGAAGTTCACGGGCAGGCCACTTTTACTACAGACACATCCTCCACACAGCCTAC	318	
QY	241	atggagatccgcagctgcacatctgaggtattctgcagttctattactgtgcaagagatact	300	
Db				
Db	319	CTGCAGCTCAGCAGCCTGACATCTGAGGACACTGCCCTCTATTATTGTAATGAGGGGACT	378	
QY	301	-----acggtcccgtttgcttactgggtccaaagggaaccaggtccacggtctctctca	351	
Db				
Db	379	CGGACTGGGCGGTACTTACTTTTGACTACTGGGGCCAAAGGACACCGGTCTCCTCTCA	438	
QY	352	ggtggagcggttcagggcgaggtggctctggcgggtggcggatcggacatcgagctcaact	411	
Db				
Db	439	GGTGGAGCGGTTTCAGGCGGAGGTGGCTCTGGCGGTGGCGGATCAGAAATGTGCTCACC	498	
QY	412	cagttctccagcaatcatgtctgcatctccagggggagaaggtcaccaatgcagctgcagtcgc	471	
Db				
Db	499	CAGTCTCCAGCAATCATGTCTGATCTCCAGGGGAGAAGGTCCACATPAACCTGCAGTGCC	558	
QY	472	agctcaagataaagttacatgcactggtacacacaaagcctgtcacctcccccaaaaga	531	
Db				
Db	559	AGCTCAAGTGTAGTTACATGTCAGTGGTTCACGAGAAAGCCAGGCATCTTCCTCAAACTC	618	
QY	532	tgatttatgacacatccaaactggcttctggagtcctgctgcgttcagtgagcagtcggg	591	


```

/strain="BALB/C"
/db_xref="taxon:10090"
/tissue_type="SPLENCYTES"
1. .732
/note="unnamed protein product; protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA02927.1"
/db_xref="PID:e306392"
/db_xref="PID:g2300754"
/db_xref="GI:2300754"
/translation="EVQLQSGAEELVKPGASVKLSCKASGYTFTSHWMHWVKQRAGQ
LEWIGFENGRGNTISNTEKRAITVDKSSATYQMQLSLTSEDSAVYICASRDYD
YDRFYDWQGTQITVSSGGGGGGGGSDIELTQSLPAIMSGASPGKEVTMTCS
ASSSVYMWYQQKPSPPRLLIYDINSLASGVVPRFSGSGSGTYSLTISRMEADA
ATVYQQQWSYYPWMYIPGGTKLEIK"
172 a 196 c 209 g 155 t
BASE COUNT
ORIGIN

```

	Very Match	70.4%	Score 504.6;	DB 5;	Length 732;	
	Best Local Similarity	83.7%;	Pred. No. 2.4e-145;			
	Matches 612;	Conservative 0;	Mismatches 101;	Indels 18;	Gaps 3;	
<hr/>						
QY	2	aggtgaactcgcagctcaggaaacctgaactggatggnacccctggggcttcagtgaagatat	61			
Db						
2	AGGTGCA3CTCAGCAGCTCGGGGTGAACCTGGTGAAGCCCTGGSGCTTCAGTGAAGTTGT	61				
QY	62	cctcgagactcttgganacaattcaactgaatacacaccatgcactgggtgaagcagagcc	121			
Db						
62	CCTCAAAGGCTTCGGGTATACCCTCACGCCACTGGATGCACCTGGGTGAAGCAGAGGG	121				
QY	122	atgcaagagaccttgagttggattgaggtattaatcctaacaatggtgtaactaacata	181			
Db						
122	CTGACAAAGCCCTTGAGTGGATCGGAGAGTTTAATCCAGAACGGCGGTACTAACAATA	181				
QY	182	agcgaagtccaaggccaagccatctgactgtagacaagctctccagcacagccata	241			
Db						
182	ATGAGAAATTCAAGCAAGGCCACACTGACTGTAGACAAATCTCTCCAGCACAGCCTACA	241				
QY	242	tggagctccgcagcctgacatctgagattctgcagctctattactgtgcaag-agatact	300			
Db						
242	TGCAACTCAGCAGCCTGACATCTGAGGACTCTGCGGTCTATTACTGTGCCAGTCGGGACT	301				
QY	301	acggtccc-----tttctactgggtccaaaggacacaggtcacccgtctcct	349			
Db						
302	ATGATTACGAGCGACGGTACTTTGACTACTGGGCCCCAAGGACACCGGTCAACCTCTCCT	361				
350	caggtgagggcgggttcagggcgaggtggtctggcgtggcgatcgagacatcagatca	409				
Db						
362	CAGGTGGCGGTGGCTCGGGCGGTGGTGGGTGGGTGGCGGGGATCTGCATTCAGATCA	421				
QY	410	ctcagttccagcaaatcatgtctgcattccaggggagaaggtcacatgacctgcagtg	469			
Db						
422	CCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAGAAGTCAACCATGACCTGCAGTG	481				
QY	470	gcagctcaagtataaagttacatgcactgggtaccagcagaagcctgtcacctcccccaaaa	529			
Db						
482	CCAGCTCAAGTGTAAATTACATGTACTGGTACACAGCAGAAAACAGGATCTCCCCCCAGAC	541				
QY	530	gatggatttatgacacatcaaactcgctctggagtccctgtcgtttcagtgggcagtg	589			
Db						
542	TCCTGATTTATGACACATCCAACCTGGCTTCGGAGTCCCTGTTCCGTTTCAGTGGCAGTG	601				
QY	590	ggctgggaacctcttatctctcaaatcagacatggaggctgtgatgtagtctgccactt	649			
Db						
602	GGTCTGGGACCTCTTACTCTCTCAAAATCAGCCGAATGAGGCTGAAGATGCTGCCACTT	661				
QY	650	attactgccatcagcgagtagttaacc-----gtcacagttcgggtgcgtgggacacagt	703			
Db						
662	ATTACTCCAGCAGCTGAGTAGTTTACCACCCCAATGTACACGTTCGGAGGGGGACCAAGC	721				
QY	704	tggaaataaaaa	714			

